

### DESCRIPCIÓN BREVE

Este documento trata sobre el análisis de los principales sitios de documentación disponibles en R para graficar

Luis Eduardo Bahena Castillo

9°C IDyGS

T501

Ing. Maximiliano Carsi Castrejón – Extracción y Conocimiento en Bases de Datos



# INTRODUCCIÓN

Analizar los principales sitios de documentación disponible en R para graficar y para cada sitio escoger 3 gráficos de interés y justificar

- <https://ggplot2.tidyverse.org/index.html>
- <https://r-charts.com/miscellaneous/ggsoccer/>
- <https://plotly.com/r/>
- <https://r-graph-gallery.com>

## DESARROLLO

### Introducción

La visualización de datos es una parte esencial del análisis estadístico, y el lenguaje de programación R es reconocido por su capacidad robusta para crear gráficos sofisticados y personalizados. Con una gran variedad de paquetes y bibliotecas disponibles, R ofrece múltiples enfoques para la creación de gráficos, cada uno con sus propias fortalezas y áreas de aplicación. En este informe, se analizarán cuatro sitios principales de documentación y recursos para la visualización en R: **ggplot2**, **R-Charts**, **Plotly**, y **R Graph Gallery**. Para cada sitio, se seleccionarán tres gráficos de interés, justificando su importancia y aplicación en el contexto del análisis de datos.

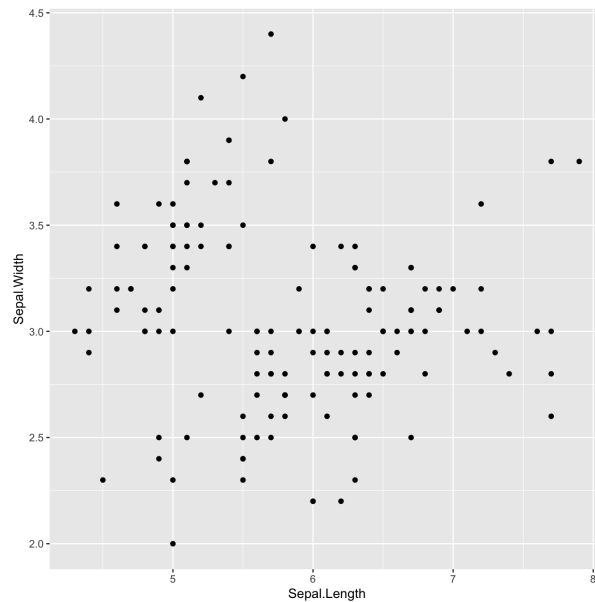
### 1. ggplot2 (Tidyverse)

El sitio oficial de **ggplot2** es parte del ecosistema Tidyverse, un conjunto de paquetes en R que facilita el manejo y análisis de datos. **ggplot2** se ha convertido en una herramienta fundamental para la creación de gráficos en R debido a su enfoque basado en "gramática de gráficos", lo que permite una gran flexibilidad y personalización.

#### Gráfico 1: Scatter Plot (Diagrama de dispersión)

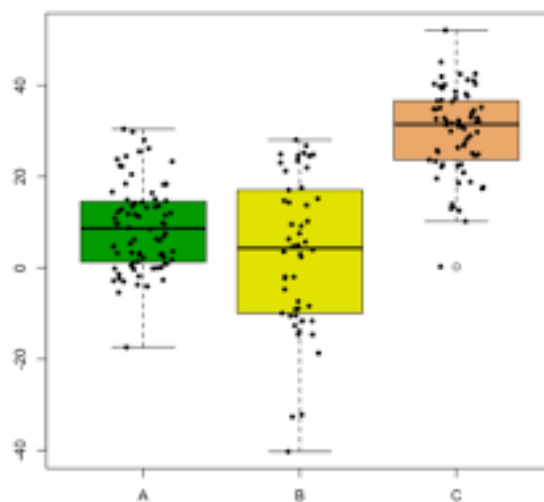
**Justificación:** El diagrama de dispersión es uno de los gráficos más utilizados en el análisis de datos. Es especialmente útil para visualizar la relación entre dos variables continuas, detectando patrones, tendencias y posibles outliers. En **ggplot2**, se pueden

agregar capas adicionales como líneas de tendencia y áreas de sombreado para enriquecer la interpretación de los datos.



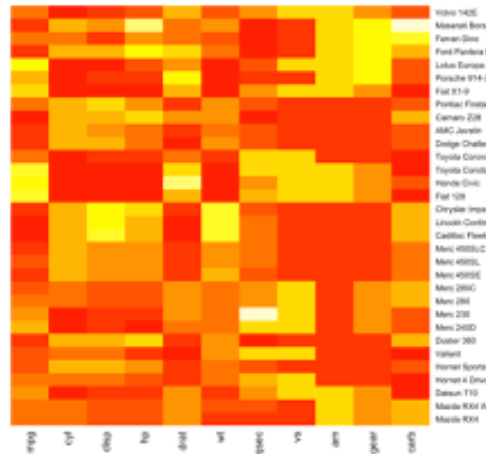
## Gráfico 2: Box Plot (Diagrama de caja)

**Justificación:** El box plot es crucial para resumir la distribución de una variable continua, mostrando la mediana, los cuartiles y posibles valores atípicos. En análisis exploratorio de datos (EDA), los box plots son herramientas clave para comparar distribuciones entre diferentes grupos categóricos.



### Gráfico 3: Heatmap (Mapa de calor)

**Justificación:** Los heatmaps son ideales para visualizar matrices de datos, como correlaciones entre variables o patrones en datos de alta dimensión. En **ggplot2**, los heatmaps pueden ser personalizados con diferentes esquemas de colores y anotaciones para resaltar áreas de interés.

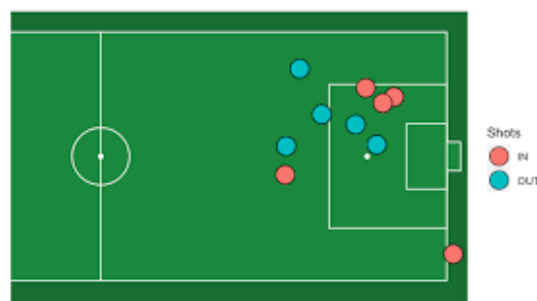


## 2. R-Charts (ggsoccer)

**R-Charts** es un recurso en línea que proporciona una amplia gama de gráficos específicos para diversas áreas. Uno de los subdominios más interesantes es **ggsoccer**, que se especializa en la visualización de datos relacionados con el fútbol.

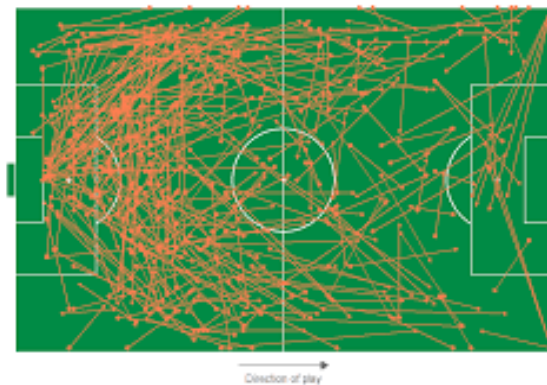
### Gráfico 1: Soccer Field Plot (Gráfico de campo de fútbol)

**Justificación:** Este gráfico es esencial para analizar y visualizar eventos en un partido de fútbol, como tiros a gol, posiciones de jugadores, y pases. El uso de un campo de fútbol como fondo permite una representación clara y contextualizada de los datos.



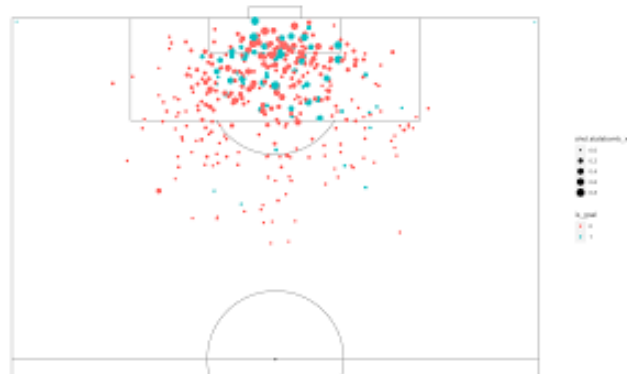
## Gráfico 2: Pass Map (Mapa de pases)

**Justificación:** El mapa de pases es una herramienta poderosa para visualizar las combinaciones de pases entre jugadores durante un partido. Este gráfico ayuda a identificar patrones de juego, fortalezas y debilidades en la distribución del balón.



## Gráfico 3: Shot Map (Mapa de tiros)

**Justificación:** El mapa de tiros muestra la ubicación y el resultado de cada intento de gol. Es crucial para analizar la eficacia de los jugadores y las tácticas de un equipo, proporcionando insights sobre áreas de mejora.

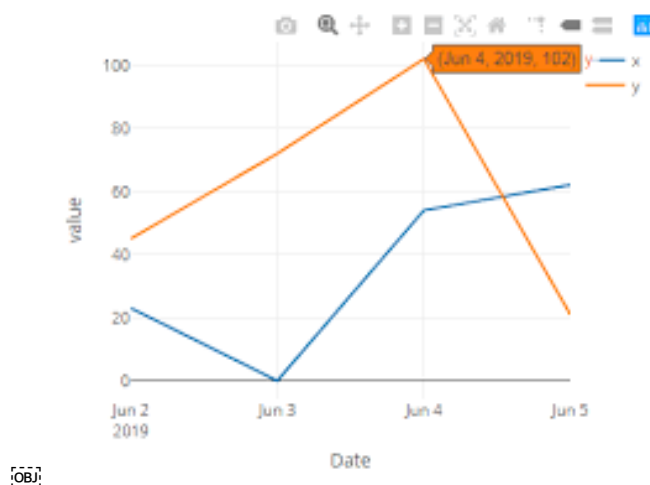


## 3. Plotly

**Plotly** es una biblioteca de visualización interactiva que se integra con R. Es particularmente útil cuando se necesita crear gráficos interactivos que permiten una exploración más profunda de los datos por parte del usuario final.

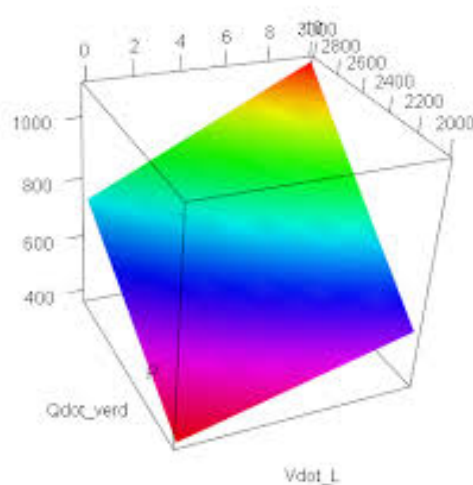
## Gráfico 1: Interactive Line Plot (Gráfico de línea interactivo)

**Justificación:** Los gráficos de línea son esenciales para visualizar tendencias a lo largo del tiempo. **Plotly** añade una capa de interactividad, permitiendo al usuario explorar diferentes puntos de datos, hacer zoom en áreas específicas y observar valores exactos, lo que mejora la comprensión y análisis de series temporales.



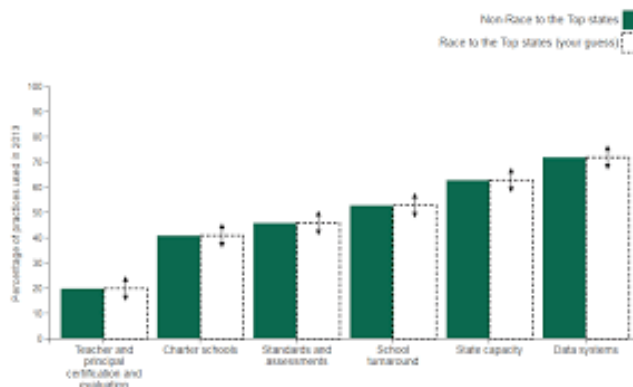
## Gráfico 2: 3D Surface Plot (Gráfico de superficie 3D)

**Justificación:** El gráfico de superficie 3D es indispensable para representar relaciones complejas entre tres variables continuas. En **Plotly**, este tipo de gráfico es altamente interactivo, lo que permite al usuario rotar, acercar y alejar la vista para comprender mejor la topología de los datos.



### Gráfico 3: Interactive Bar Chart (Gráfico de barras interactivo)

**Justificación:** Los gráficos de barras son comunes para comparar cantidades entre diferentes categorías. La interactividad proporcionada por **Plotly** permite que los usuarios seleccionen o deselectionen categorías, vean datos exactos y descubran más detalles sobre cada categoría al interactuar con el gráfico.



## 4. R Graph Gallery

**R Graph Gallery** es un repositorio de gráficos creados en R, que ofrece ejemplos prácticos de una amplia variedad de tipos de gráficos, desde los más básicos hasta los más complejos y personalizados.

[OBJ]

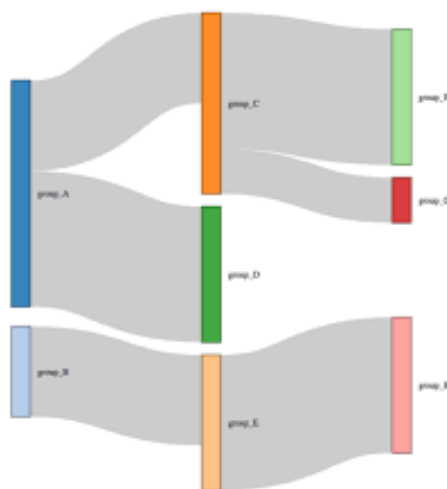
### Gráfico 1: Circular Barplot (Gráfico de barras circular)

**Justificación:** El gráfico de barras circular es una variación estética del gráfico de barras tradicional, donde las barras se disponen en forma circular. Es especialmente útil para visualizar datos categóricos con un gran número de categorías, manteniendo una presentación compacta y visualmente atractiva.



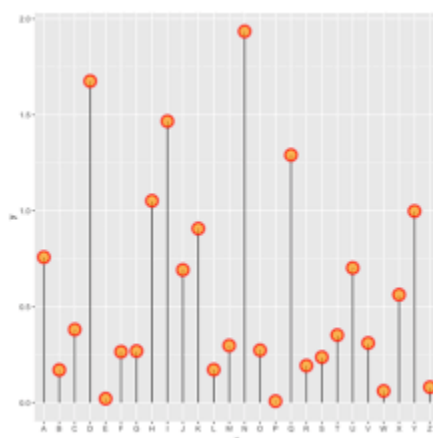
## Gráfico 2: Sankey Diagram (Diagrama de Sankey)

**Justificación:** El diagrama de Sankey es crucial para visualizar flujos y conexiones entre diferentes entidades. Es ampliamente utilizado en análisis de energía, migración de datos y análisis de procesos para mostrar cómo se distribuyen los flujos a través de diferentes nodos.



## Gráfico 3: Lollipop Chart (Gráfico de piruletas)

**Justificación:** El gráfico de piruletas es una alternativa al gráfico de barras que puede ser más eficaz para representar comparaciones cuando se desea evitar la congestión visual. Es particularmente útil cuando se tiene un gran número de categorías, permitiendo una visualización clara y precisa.





# CONCLUSIÓN

Los sitios de documentación y recursos analizados ofrecen una amplia gama de opciones para la visualización de datos en R, reflejando la versatilidad y potencia del lenguaje para atender diversas necesidades en el análisis de datos. Cada uno de estos sitios se destaca por enfoques y especialidades únicas, proporcionando herramientas que se adaptan a diferentes contextos y tipos de datos.

**ggplot2**, como parte del ecosistema Tidyverse, se erige como una herramienta esencial para aquellos que buscan flexibilidad y capacidad de personalización en sus gráficos. La "gramática de gráficos" que propone **ggplot2** permite construir visualizaciones complejas de manera sistemática y coherente, haciendo posible no solo la creación de gráficos básicos, sino también de representaciones sofisticadas que se ajustan a las necesidades específicas del análisis. Esta flexibilidad convierte a **ggplot2** en una opción predilecta tanto para análisis exploratorio como para la presentación final de resultados.

Por otro lado, **R-Charts** (específicamente su subdominio **ggsoccer**) destaca por su enfoque especializado en la visualización de datos relacionados con el fútbol. Esta especialización es un claro ejemplo de cómo R puede adaptarse a nichos específicos, ofreciendo soluciones visuales que son extremadamente relevantes para analistas deportivos y entrenadores que buscan comprender mejor el comportamiento del juego. Los gráficos como los mapas de pases y de tiros no solo son intuitivos y fáciles de interpretar, sino que también proporcionan un análisis profundo y detallado de las dinámicas dentro del campo de juego, lo que puede ser crucial para la toma de decisiones tácticas.

**Plotly** introduce un elemento crucial en la visualización de datos: la interactividad. En un mundo donde la exploración de datos en tiempo real es cada vez más importante, **Plotly** ofrece una plataforma donde los usuarios pueden interactuar directamente con los gráficos, permitiendo una comprensión más profunda de los datos a través de la manipulación directa de las visualizaciones. La posibilidad de crear gráficos tridimensionales, mapas de calor interactivos y gráficos de líneas dinámicos hace que **Plotly** sea una herramienta invaluable para analistas que requieren un nivel adicional de detalle y personalización en la exploración de sus datos.

Finalmente, **R Graph Gallery** se presenta como un repositorio extenso y variado, donde los usuarios pueden encontrar ejemplos prácticos de prácticamente cualquier tipo de gráfico que se pueda necesitar en R. La diversidad de gráficos disponibles no solo sirve como inspiración para los usuarios, sino que también proporciona soluciones listas para implementar en diferentes contextos analíticos. Desde gráficos tradicionales como barras y líneas, hasta visualizaciones más creativas como los diagramas de Sankey o gráficos de piruletas, **R Graph Gallery** es un recurso indispensable para cualquier analista de datos que busca expandir su repertorio de herramientas visuales.