

Visualización estática con R: ggplot2

Manuel Rueda
mrueda@afi.es

Abril de 2021

Índice

1. Fundamentos de visualización de datos
2. ggplot2
3. Materiales
4. Anexo: Ampliación de conceptos (Visualización estática con Python)

1 | *Fundamentos de visualización de datos*

Visualización de datos

- Técnicas y métodos que transmiten el significado de un conjunto de datos mediante su representación en un contexto visual.
- Técnicas y métodos que permiten unir las disciplinas del diseño y la estadística para facilitar la extracción de información y comprensión de un conjunto de datos.
- Técnicas y métodos que permiten **comunicar información de una forma visual, clara, eficiente y efectiva**, de modo que se permita la comprensión y análisis de los datos subyacentes a los elementos gráficos e, incluso, faciliten el razonamiento más allá de los mismos.

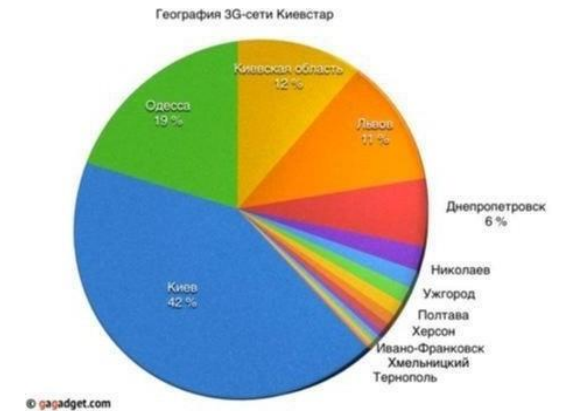
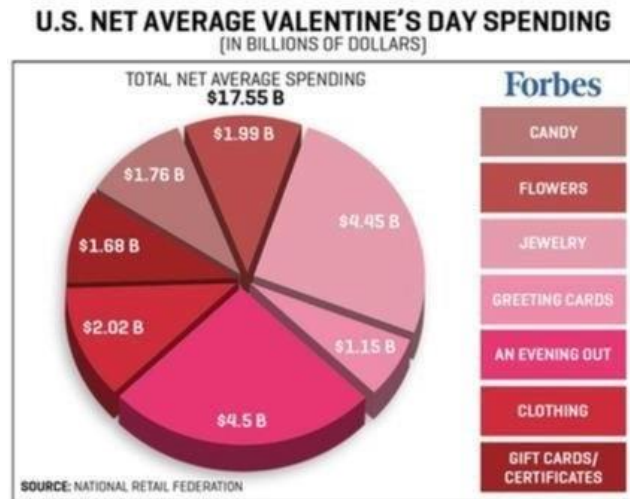
Buenas visualizaciones (Tufte, 1983)

- Son rigurosas con los datos presentados.
- Aprovechan al máximo la información contenida en el conjunto de datos disponible.
- Tienen claro el objetivo perseguido, es decir, no confunden a quien las consume con información innecesaria o difícilmente interpretable.
- Tienen clara la audiencia a quien están dirigidas.

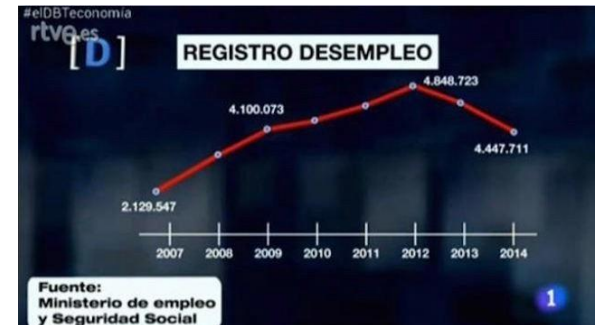
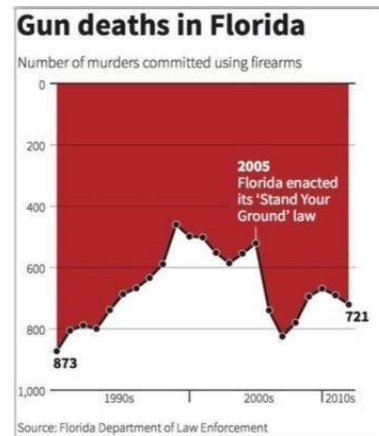
Tipos de visualización

	Exploratoria	Divulgativa
Objetivo	Entender	Comunicar
Audiencia objetivo	Analistas	Terceros
Tipo de consumo	Interactivo	Estático
Tiempo de elaboración	Corto	Alto
Calidad del diseño	Baja	Alta
Volumen de conclusiones	Poco	Mucho

Ejemplos de mala visualización de datos



Ejemplos de mala visualización de datos



2 | ggplot2

¿Qué es ggplot2?

- Es un paquete de R que permite la creación de visualizaciones 2D basadas en datos.
- Al contrario que otros paquetes de gráficos (p.e. base, lattice...) no incluye funciones para la creación de gráficos “predefinidos”.
- Entiende los gráficos como objetos formados por componentes asociados a diferentes elementos visuales.
- Estos componentes y su interacción están descritos en una “gramática de gráficos basada en capas” que establece los elementos básicos de un gráfico y las “normas” que deben seguir.
- Permite la creación de gráficos de forma iterativa (“por capas”).

¿Por qué ggplot2?

- Presenta buenas posibilidades de exportación.
- Es flexible.
- El diseño incremental lo hace útil en visualización exploratoria y presentación de resultados.

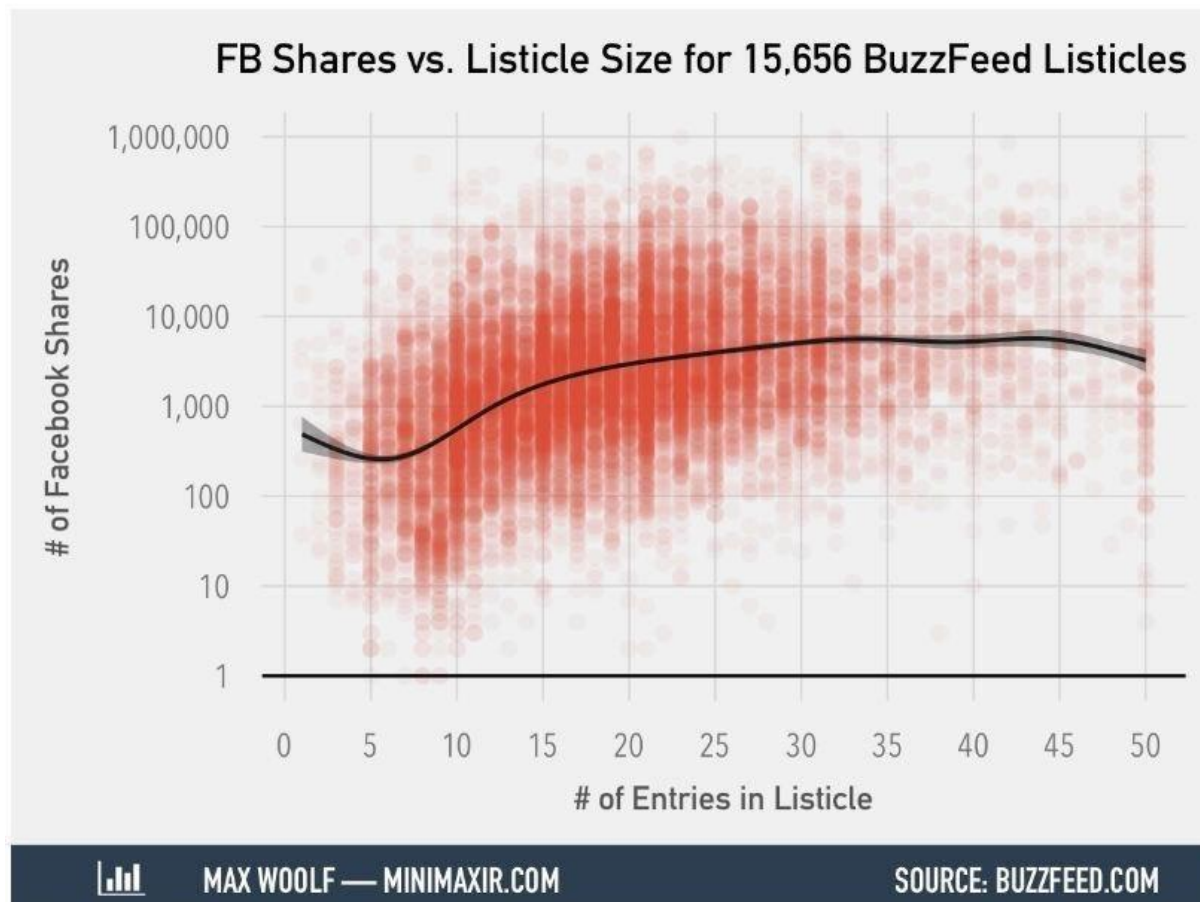
Desventaja de base / Motivación de ggplot2

- Tratamiento del gráfico como una pura imagen “estática” creada en la primera llamada.
- No hay ningún recálculo a la hora de añadir “anotaciones” o elementos adicionales al primer gráfico.
- No incluye elementos adicionales basados en datos (p.e. leyendas) sino que hay que generarlos de forma manual.
- No define un framework común para la creación de gráficos, sino que da un conjunto de funciones “predefinidas” para cada tipo de gráfico, cada una de ellas con sus parámetros y funcionamiento específico e independiente.
- En resumen, dificulta la creación de gráficos de forma iterativa y no ofrece una base común para la generación de visualizaciones.

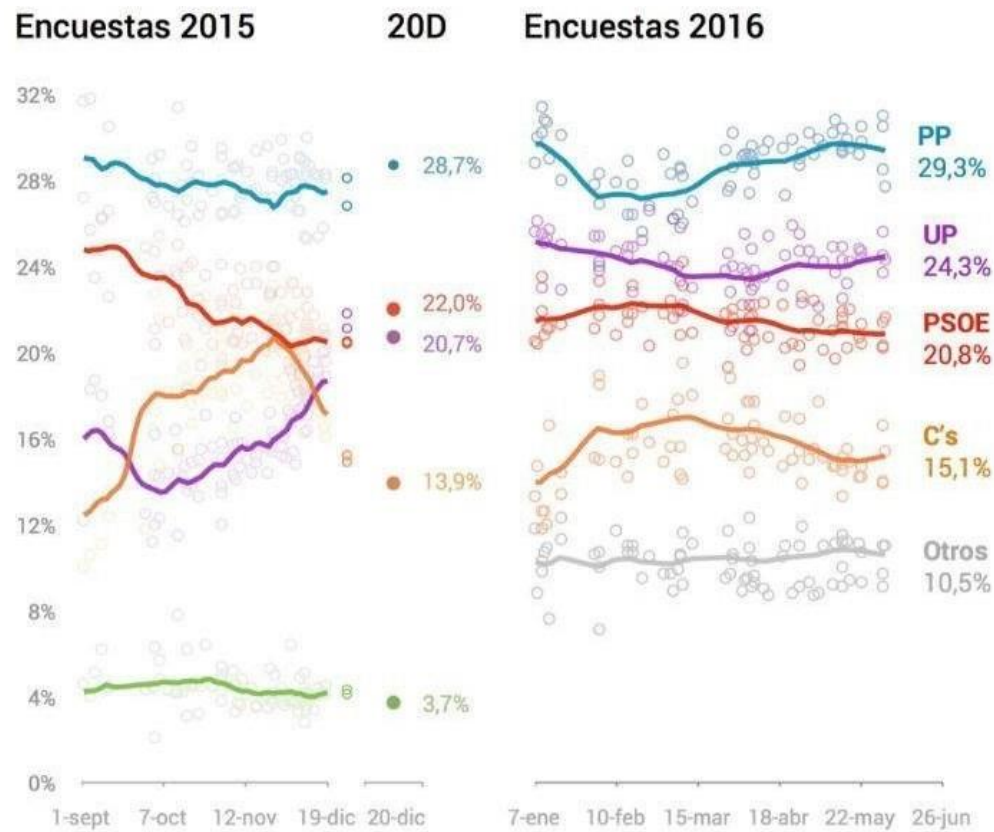
Componentes de la gramática de gráficos

Componente	Descripción	Obligatorio	Ejemplos
Data	Set de datos en que se basa el gráfico	Sí	Cualquier data.frame
Aesthetics	Características visuales asociadas a los datos	Sí	x, y, color, fill, alpha
Geometries	Elementos visuales que representan cada observación	Sí	point, bar, line
Statistics	Agregaciones	No	smooth, abline
Facets	Desagregación en múltiples gráficos	No	grid
Coordinates	Descripción del espacio donde se presenta el gráfico	No	cartesian, polar
Themes	Aspectos visuales no relacionados con los datos	No	axis.line, plot.title

Ejemplos (I)



Ejemplos (II)



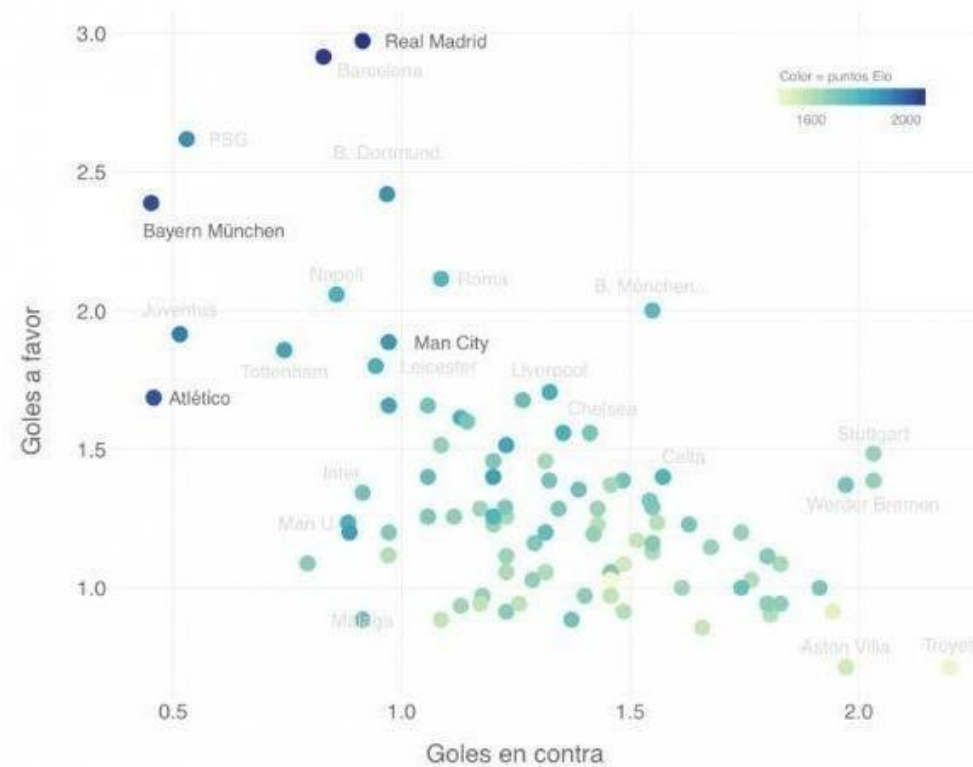
Fuente: Electograph (UP incluye confluencias).

Kiko Llaneras | EL ESPAÑOL

Ejemplos (III)

Capacidad de marcar y no encajar

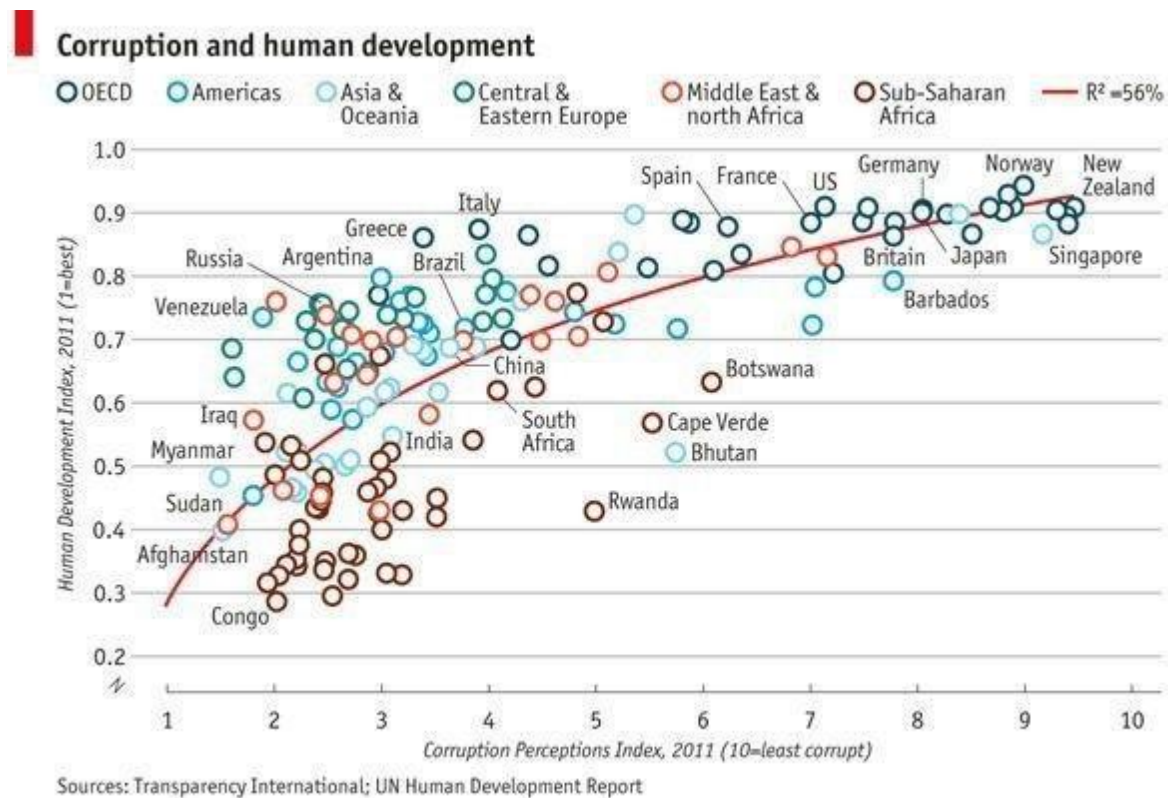
Goles en contra (x) y goles a favor (y)



Fuente: Opta Sports, Clubele (mayo 2018)

Carmona y Llaneras | JOT DOWN

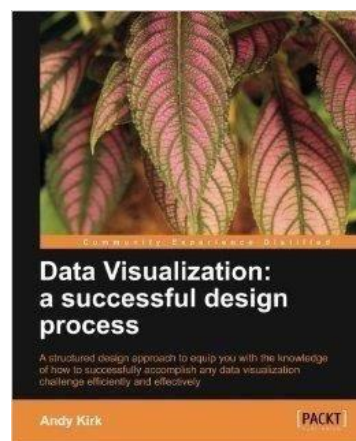
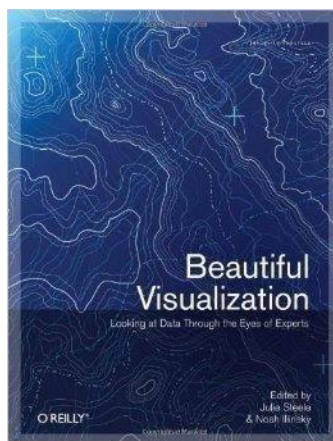
Ejemplos (IV)



3 | Materiales

Materiales sobre visualización

Data Visualization: a successful design process
Beautiful Visualization



Online:

<https://ggplot2.tidyverse.org/index.html>

<http://sape.inf.usi.ch/quick-reference/ggplot2>

<https://github.com/rstudio/cheatsheets/raw/master/data-visualization-2.1.pdf>



© 2021 Afi Escuela de Finanzas. Todos los derechos reservados.