

La importancia de la visualización. Gráficos interactivos en R.

VII Seminario Análisis de datos avanzado en Ciencias de la Salud

Carmen Lancho Martín - DSLAB

15 de octubre de 2024



- 1 Análisis exploratorio de datos vs análisis descriptivo
- 2 Ejemplos de éxitos y fracasos
- 3 Gráficos interactivos

- ¿Qué es el análisis descriptivo de datos?
- ¿Qué herramientas usa?



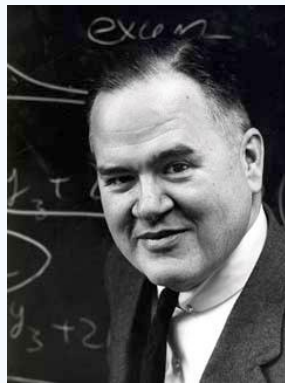
- ¿Qué es el análisis descriptivo de datos?
- ¿Qué herramientas usa?
 - Media, mediana, moda, varianza, etc.
 - Tablas de frecuencia
 - Diagramas de barras, boxplot, diagramas de dispersión, etc.

"Exploratory data analysis can never be the whole story, but nothing else can serve as the foundation stone as the first step"

"Exploratory Data Analysis is detective work"

"The simple graph has brought more information to the data analyst's mind than any other device"

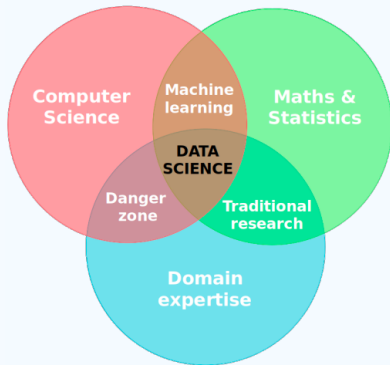
John Tukey



¿Cuándo realizo visualizaciones?



¡Siempre que sea necesario!



(a) Foundations

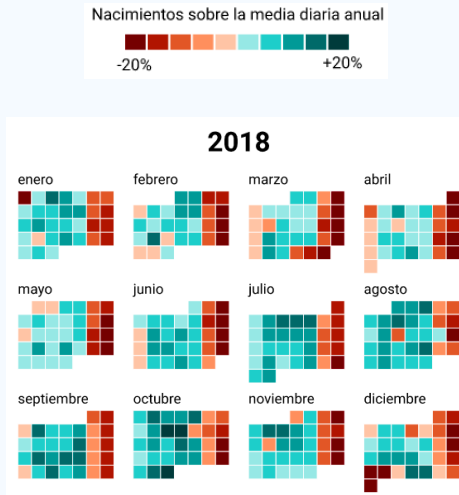


(b) Applications

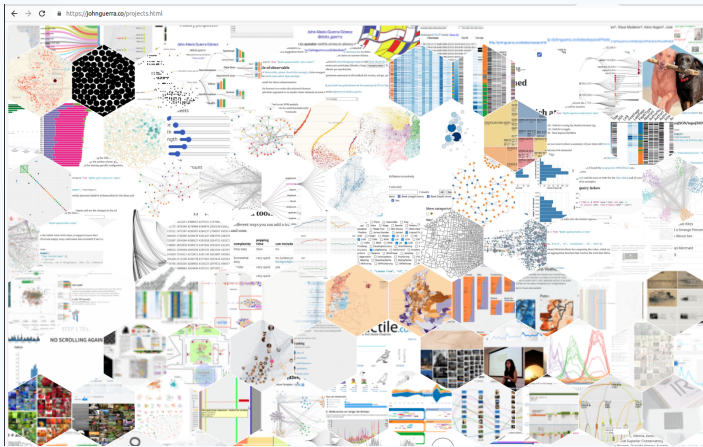
Figure 1.1: Data Science

- Entender los datos
- Descubrir patrones y comportamientos
- Explicar y transmitir la información que contienen los datos
- Transparencia e interpretabilidad
- Resumir información

Estudio natalidad: https://www.eldiario.es/nidos/no-ninos-nacen-toca-dar-luz-semana-21-probable-hacerlo-lunes-viernes_1_6400307.html

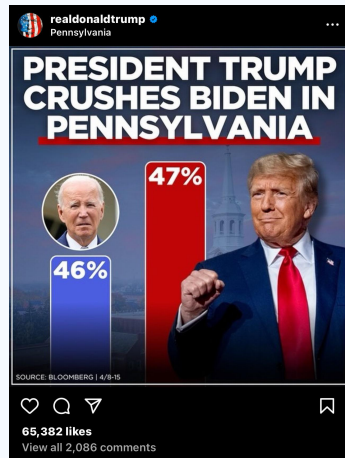
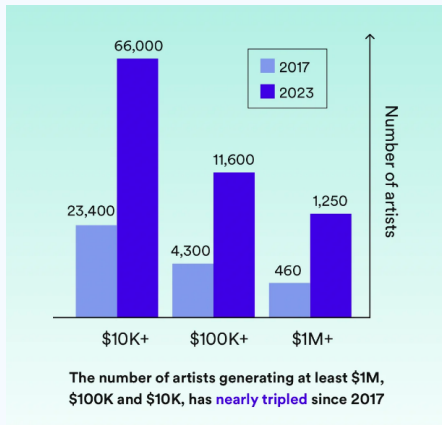


- John Alexis Guerra <https://johnguerra.co/>
- Resultados electorales Colombia:
<https://johnguerra.co/viz/resultadosSegundaVuelta2018/>



- Dashboards COVID

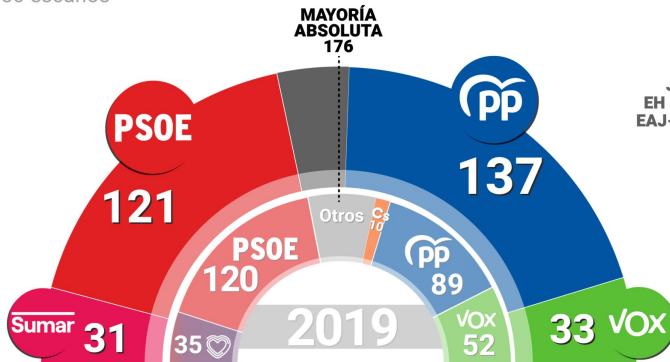
- New York Times: <https://www.nytimes.com/interactive/2021/us/covid-cases.html>
- Instituto de Salud Carlos III: <https://cnecovid.isciii.es/covid19/>
- DSLAB report
- Universidad John Hopkins, Ghana Statistical Service, etc.



Elecciones generales 2023

Congreso de los Diputados

350 escaños



Otros:

ERC - 7
Junts - 7
EH Bildu - 6
EAJ-PNV - 5
BNG - 1
CC - 1
UPN - 1

Otros en 2019: ERC (13), JuntsxCat (8), EAJ-PNV (6),
EH Bildu (5), Más País / Compromís (3), CC-NC (2),
CUP (2), NA+ (2), PRC (1), Teruel Existe (1)



Plotly Open Source Graphing Libraries

Interactive charts and maps for Python, R, Julia, Javascript, ggplot2, F#, MATLAB®, and Dash.



Plotly Python
Open Source
Graphing Library

 Star 16,126



Plotly R
Open Source
Graphing Library

 Star 2,551



Plotly Julia
Open Source
Graphing Library

 Star 423



Plotly Javascript
Open Source
Graphing Library

 Star 16,948



Plotly ggplot2
Open Source
Graphing Library

 Star 2,551



Plotly F#
Open Source
Graphing Library

 Star 654



Plotly MATLAB®
Open Source
Graphing Library

 Star 380



Plotly Dash
Open Source
Analytical App Framework

 Star 21,319

Los haremos en R usando la librería *plotly*: <https://plotly.com/r/>

Dos maneras de hacerlo:

- *ggplot* → añadimos interactividad *ggplotly*
- directamente en *plotly*

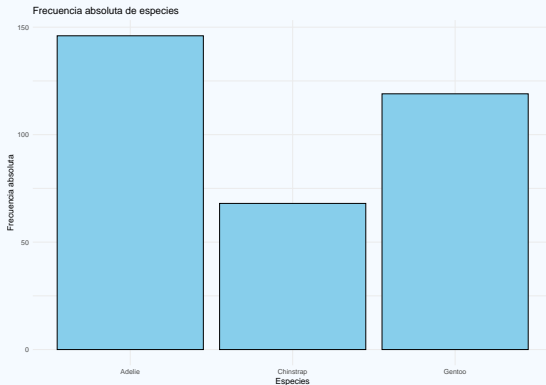
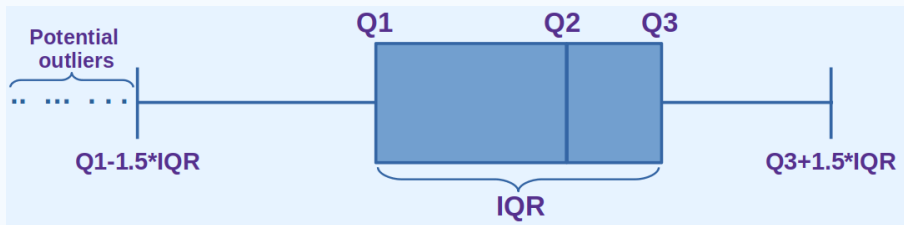


Diagrama de cajas y bigotes (boxplot)



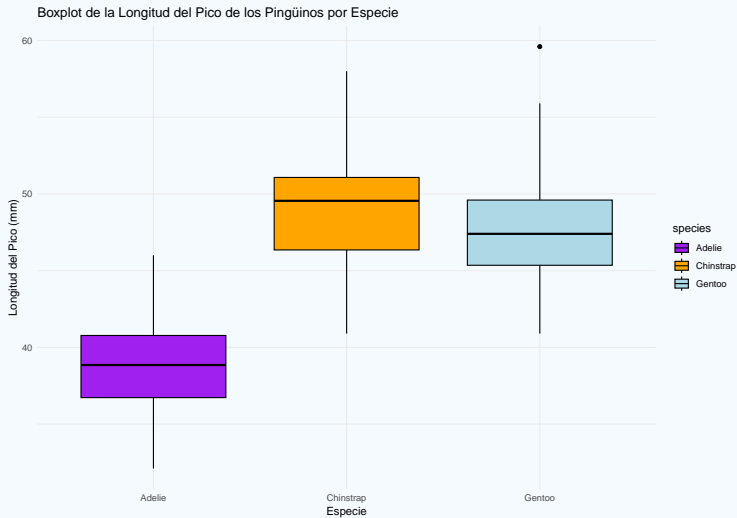
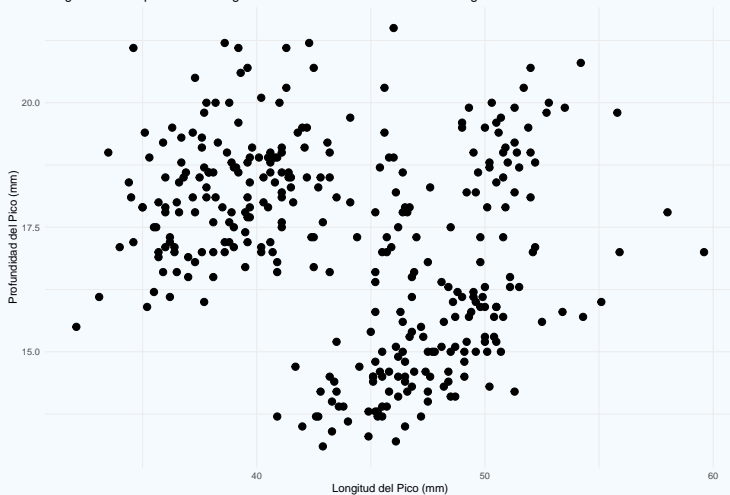
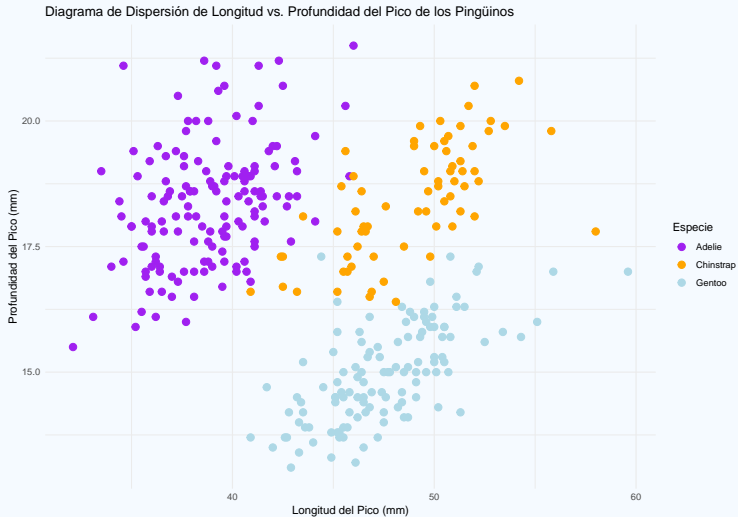


Diagrama de Dispersión de Longitud vs. Profundidad del Pico de los Pingüinos





- Modern Data Visualization with R
<https://rkabacoff.github.io/datavis/>
- <https://www.data-to-viz.com/>
- R graph gallery <https://r-graph-gallery.com/>
- R Gallery book <https://bookdown.org/content/b298e479-b1ab-49fa-b83d-a57c2b034d49/>
- ¿El mejor gráfico hecho? The Minard map
<https://bigthink.com/strange-maps/229-vital-statistics-of-a-deadly-campaign-the-minard-map/>
- <https://elartedeldato.com/>

- “Fundamentos de ciencia de datos con R” coordinado por Gema Fernández-Avilés y José-María Montero: <https://cdr-book.github.io/>
- “Estadística Aplicada a las Ciencias y la Ingeniería” escrito por Emilio L. Cano. <https://emilopezcano.github.io/estadistica-ciencias-ingenieria/index.html>
- R for Data Science: <https://r4ds.hadley.nz/eda>
 - Primera versión en castellano: <https://es.r4ds.hadley.nz/>