

Exploratory data Analysis Gráficas

Andrés Godoy
Python y Machine Learning para economistas



Agenda

1. ¿POR QUÉ? ¿POR QUÉEEEEEEEEEEEEE?
2. ¿Qué es una gráfica?
3. ¿Qué quiere mi Rey/reina?
4. ¿Qué opciones hay?

Por qué son tan importantes las gráficas



Porque evolucionaste por millones de años para interpretar miles de millones de señales lumínicas en un segundo.

Sería irónico, incluso estúpido, usar modelos de inteligencia artificial para aprender de la realidad, sin apoyarte de la mayor inteligencia para el razonamiento abstracto que existe hasta el momento: la tuya, la nuestra.

¿Por quéeeeeeee?



Procesamiento:

Nuestro cerebro está altamente adaptado para el procesamiento visual.

Las imágenes y gráficos se procesan más rápido y se retienen mejor que el texto y los números.

Patrones y Relaciones:

Las gráficas nos permiten ver patrones y relaciones entre los datos de manera intuitiva, sin tener que razonar.

Esto es clave para el aprendizaje y la comprensión.



Memoria Visual:

Tendemos a recordar información visual más fácilmente que la textual o auditiva.



«Si comunicar algo quieres, graficarlo debes.
Y ojalá bien, si pasar la materia quieres».

Haz un meme estadístico



¿Qué quieres?

¿te parece que la foto de al lado tiene que ver con el tema? No, cierto? Así es con las gráficas

1

Define el Propósito de la Gráfica:

- ¿Estás tratando de mostrar una tendencia a lo largo del tiempo?
- ¿Quieres comparar diferentes grupos o categorías?
- ¿Estás interesado en resaltar la relación entre dos o más variables?
- ¿Necesitas mostrar la distribución de un conjunto de datos?

¿Quéquieres?

2

Considera el Tipo de Datos que Tienes:

Datos Categóricos: Se refieren a nombres, etiquetas o categorías (por ejemplo, tipos de frutas, nombres de países).

Datos Numéricos: Pueden ser continuos (por ejemplo, altura, peso) o discretos (por ejemplo, número de hijos, número de coches vendidos).

Datos temporales, Datos espaciales, Datos de Texto

3

¿Cuál es tu audiencia?



«Graficar es hacer un post de Instagram, dedícale el mismo tiempo,
piensa en qué trasmites y qué dirán los demás».

Aristóteles



Desafortunadamente...

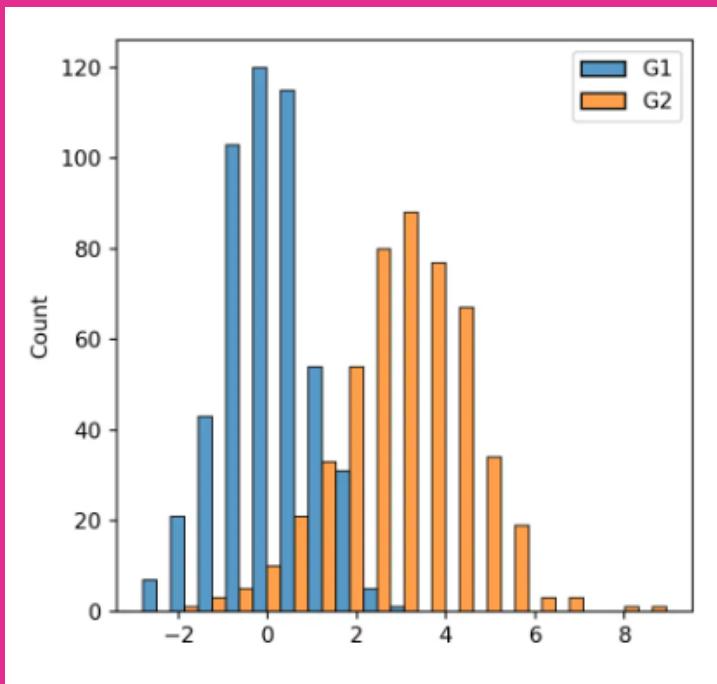
Si no sabes qué posibilidades diferentes hay y nunca has intentado imaginar algo diferente... cómo vas a terminar tu relación tóxica? Y más aún cómo vas a saber qué graficar?



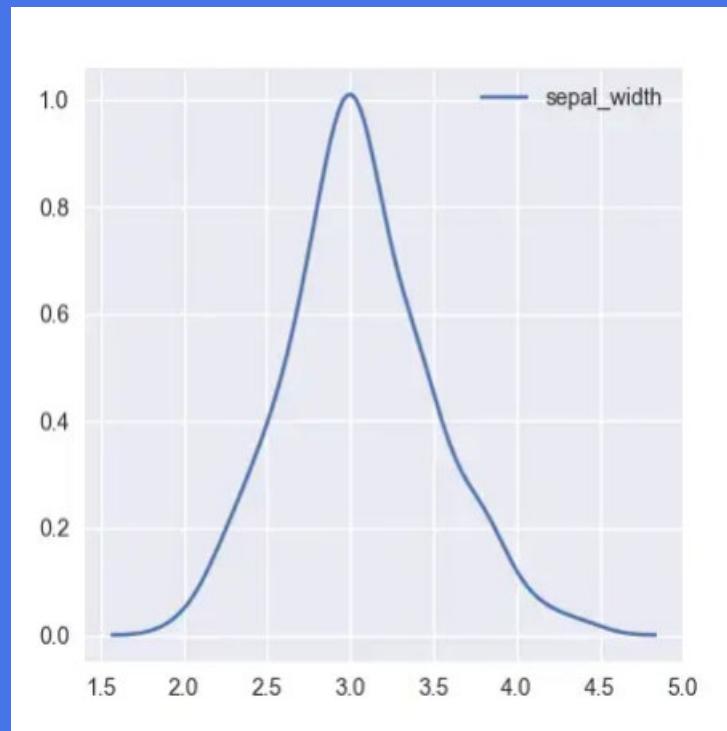
Tipos de gráficas según propósitos

- 0. Comparación entre variables categóricas
(Colorsitos y formas)
- 1. Distribución de los datos
- 2. Tendencia temporal
- 3. Correlación
- 4. Proporciones
- 5. Ranking

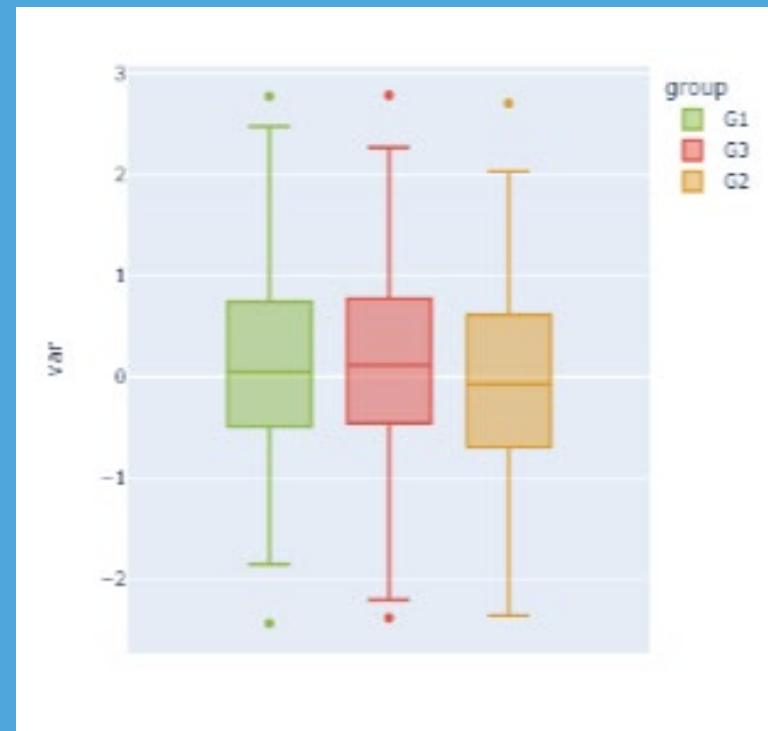
1. Distribución de los datos



Histogram

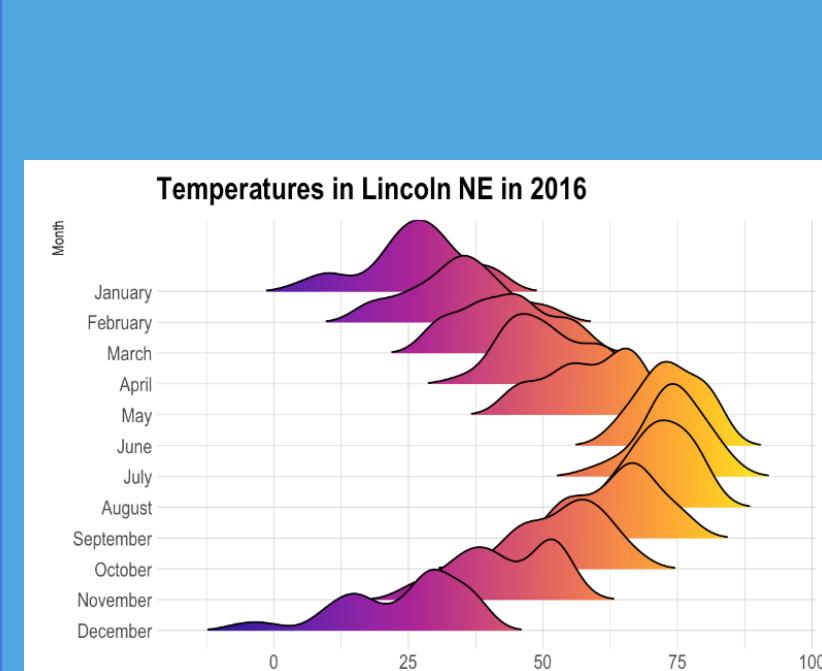
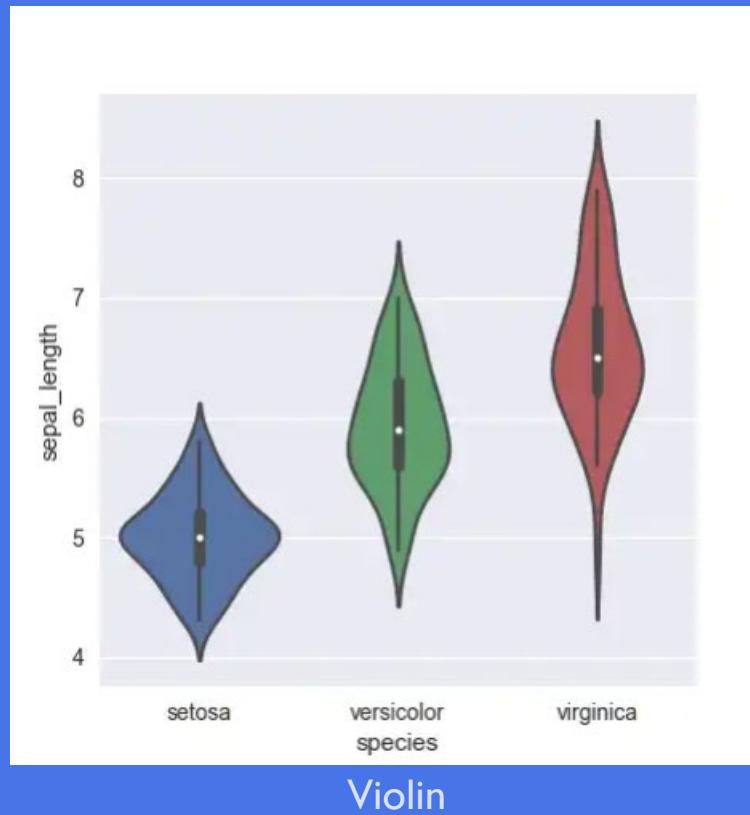
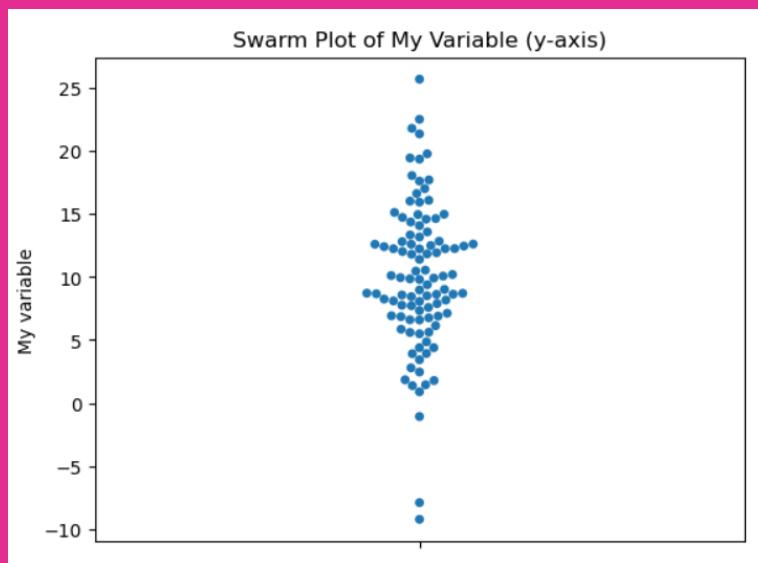


Density Graph



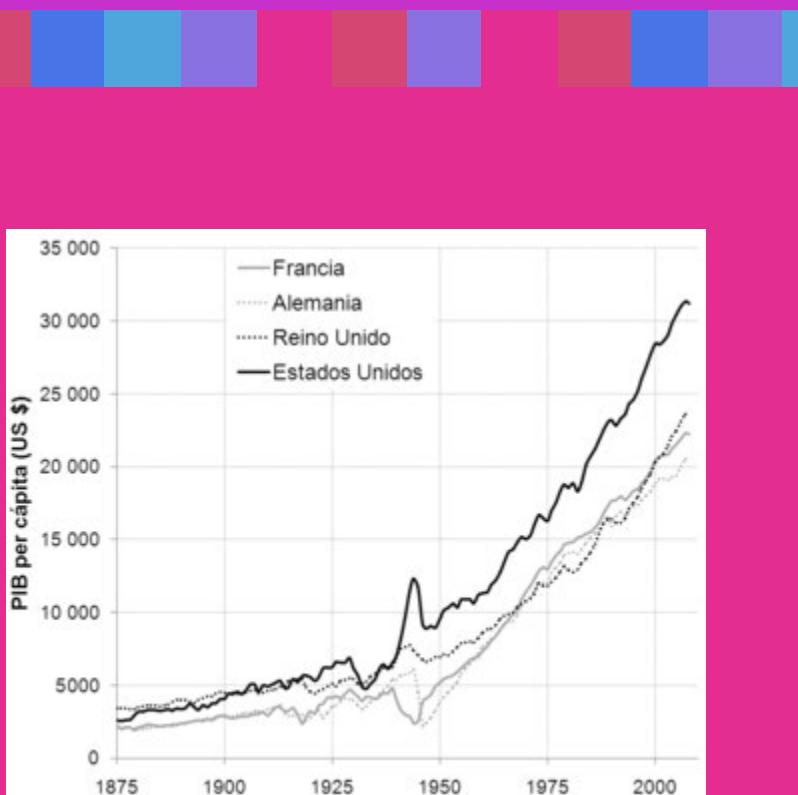
Box plot

1. Distribución de los datos



2. Temporales

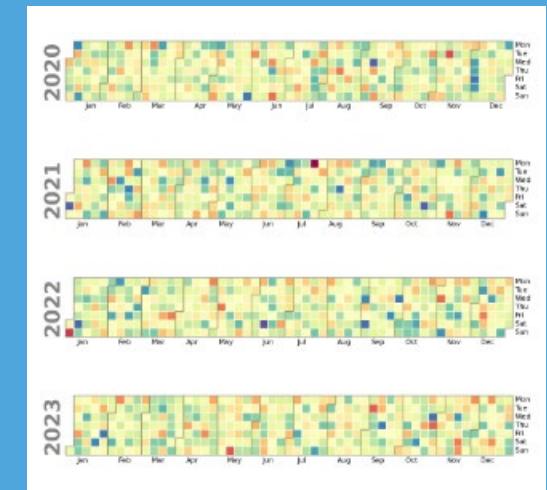
(Las líneas son normalmente para tiempo)



Line Chart



Column chart

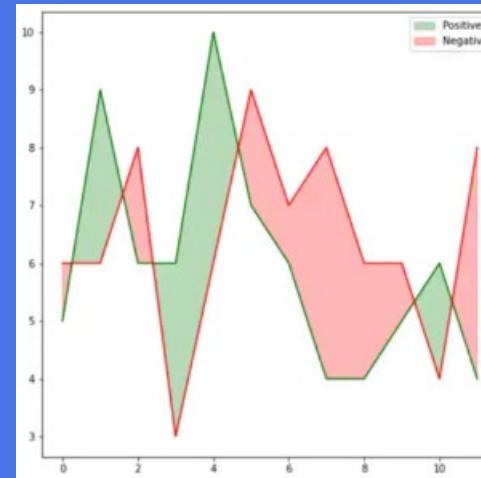


Calendar plot

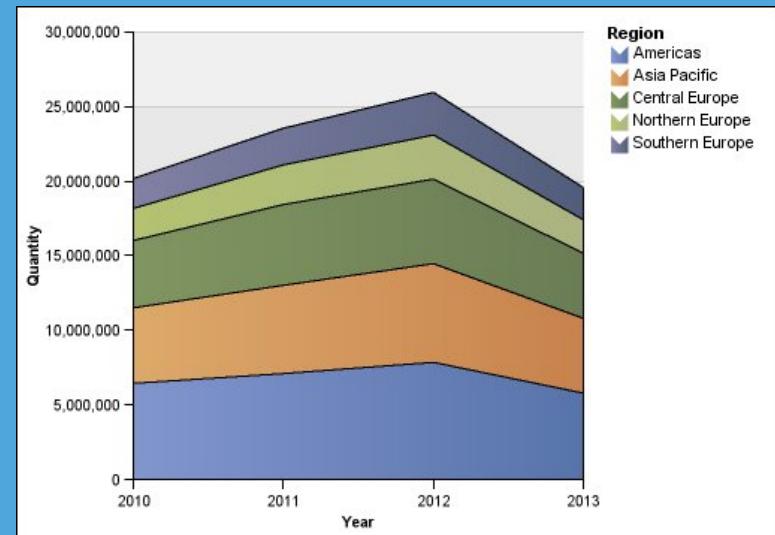
2. Temporales (las mismas líneas pero en áreas)



Area chart

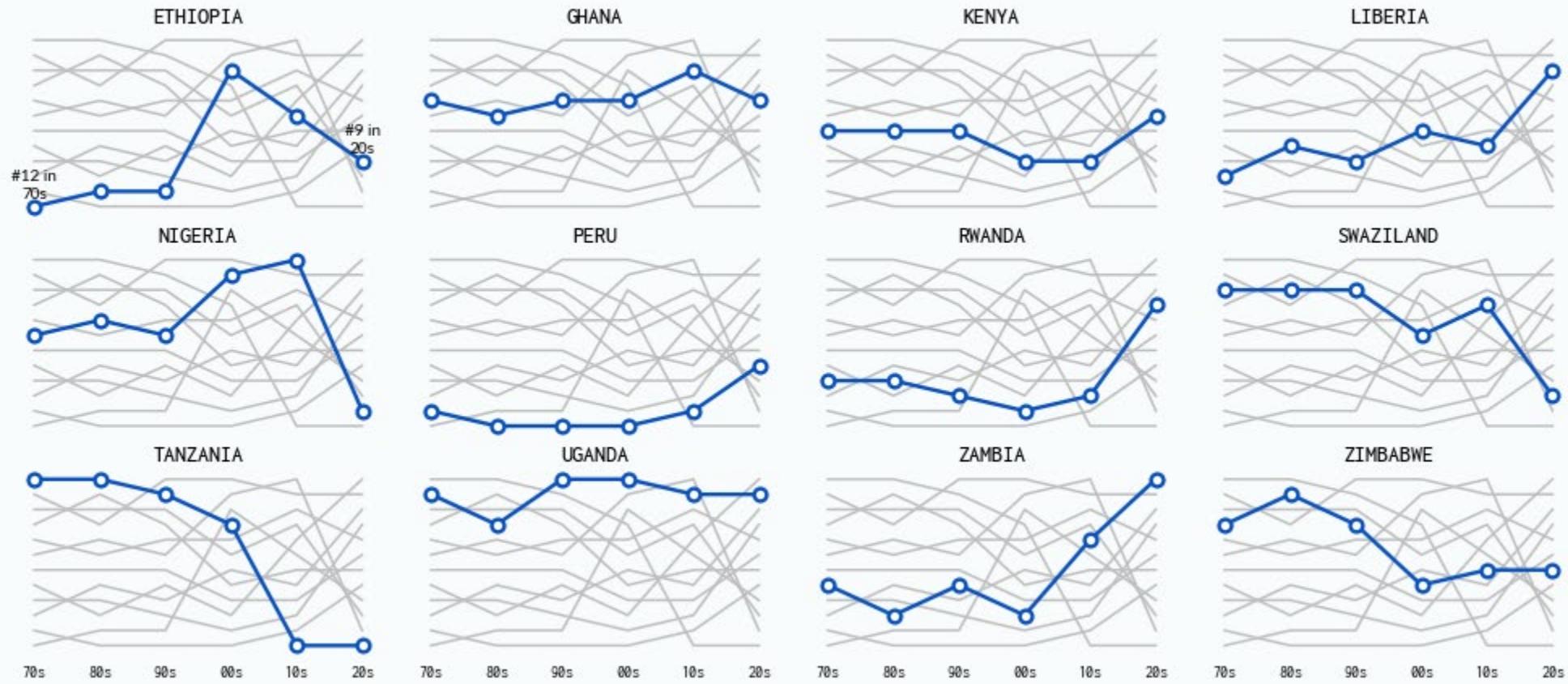


Area Chart



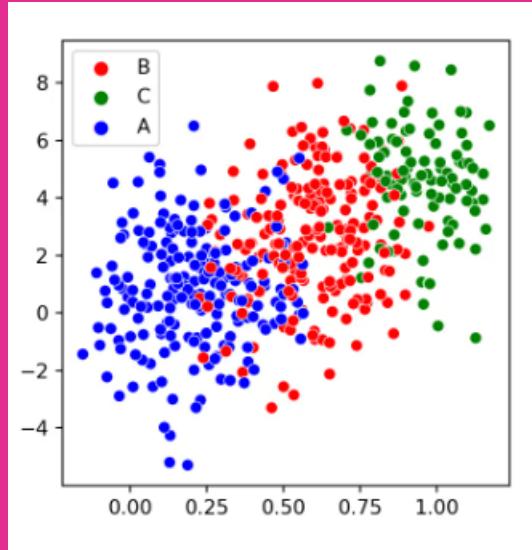
Áreas apiladas

RANKING SOME COUNTRIES BY THE NUMBER OF WATER SOURCES INSTALLATIONS BY DECADE

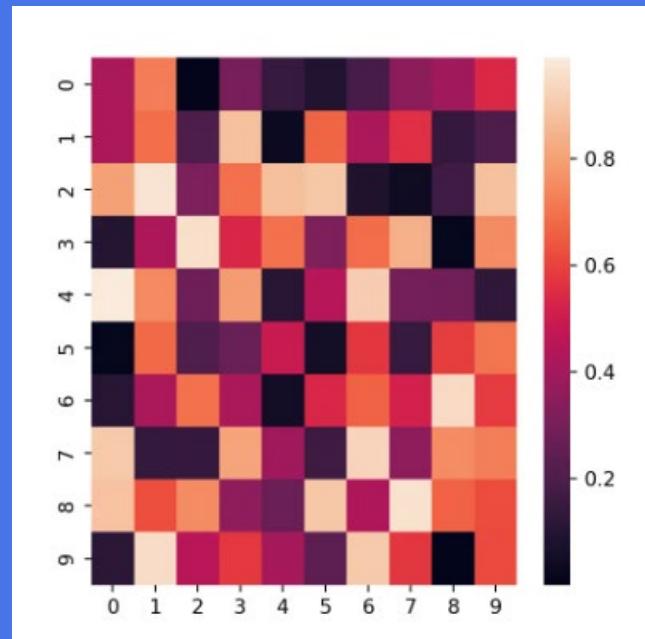


Data from Water Point Data Exchange
@issa_madjid

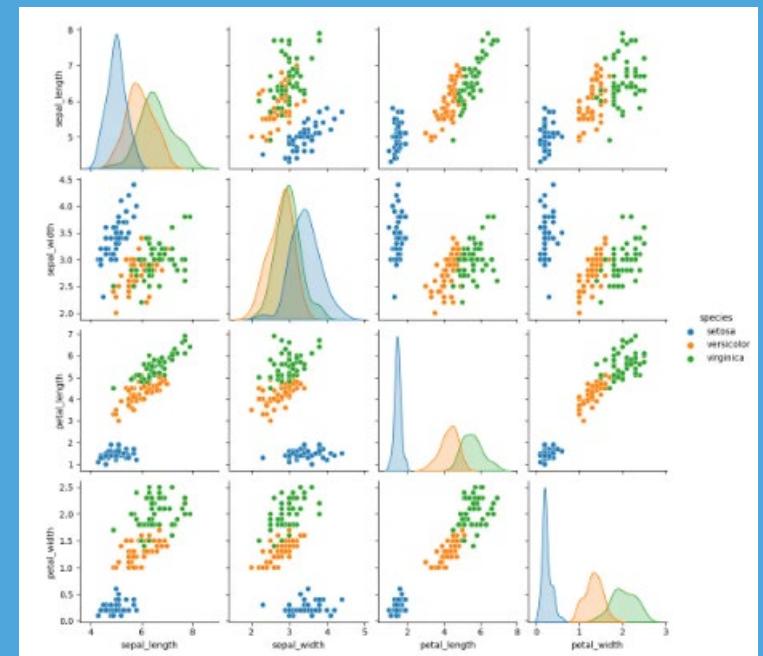
3. Correlación



Scatter plot

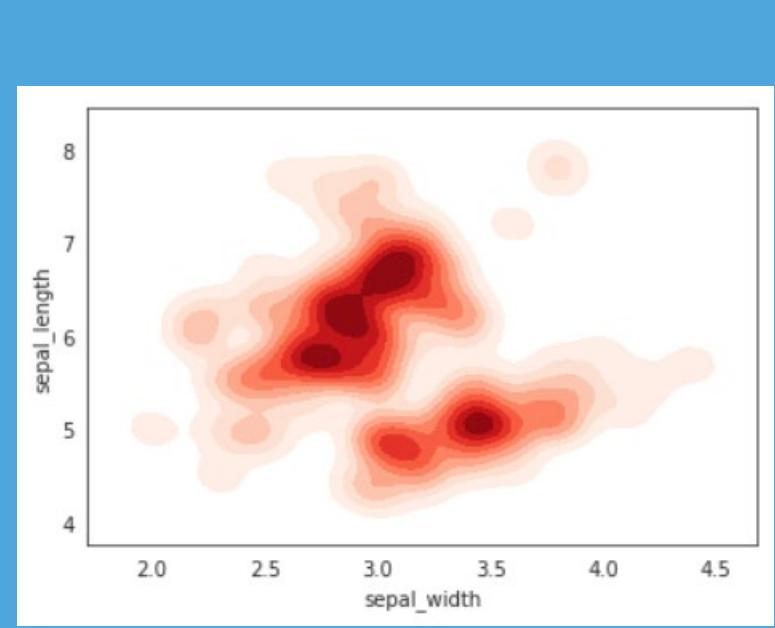
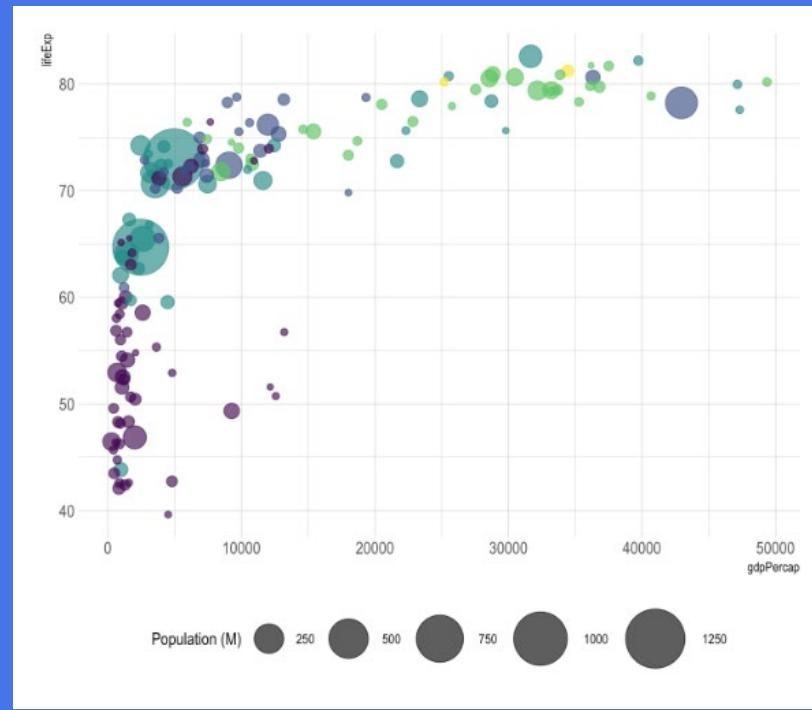
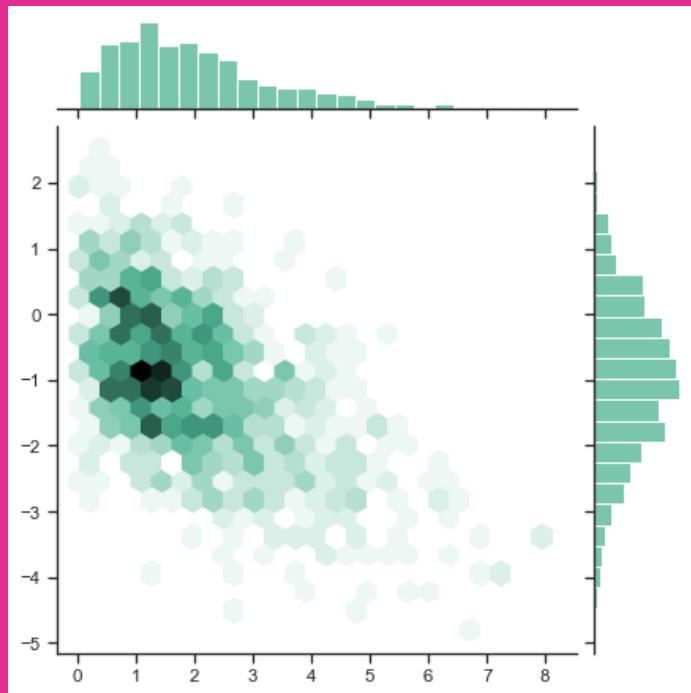


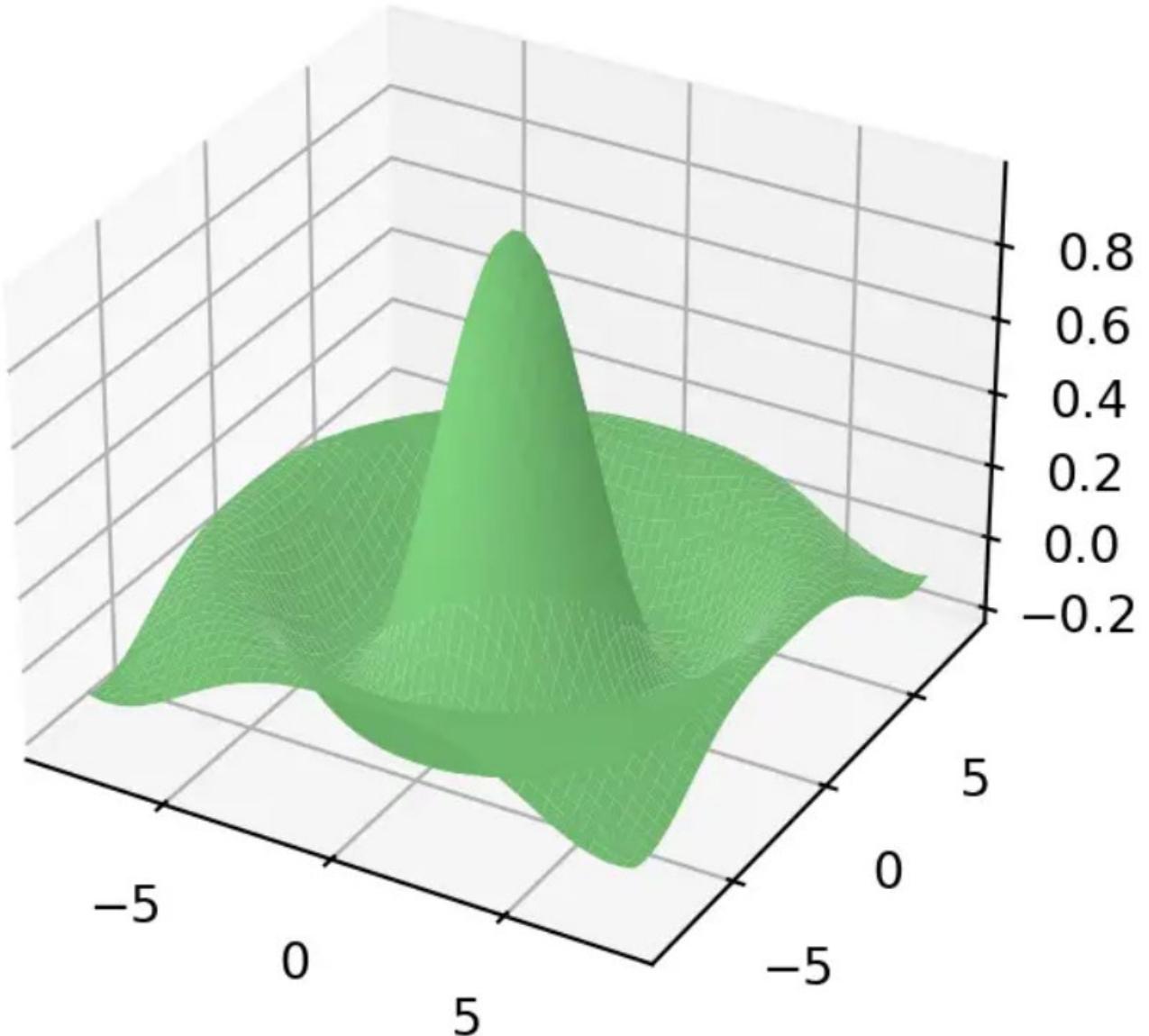
Heatmap



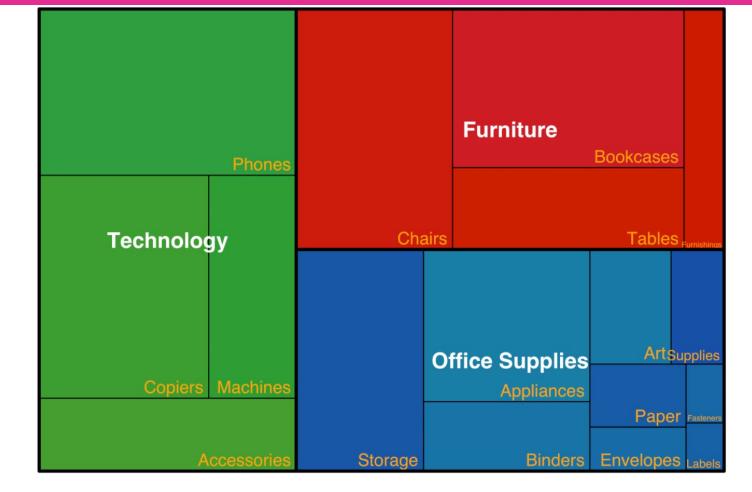
Correlograma

3. Correlación

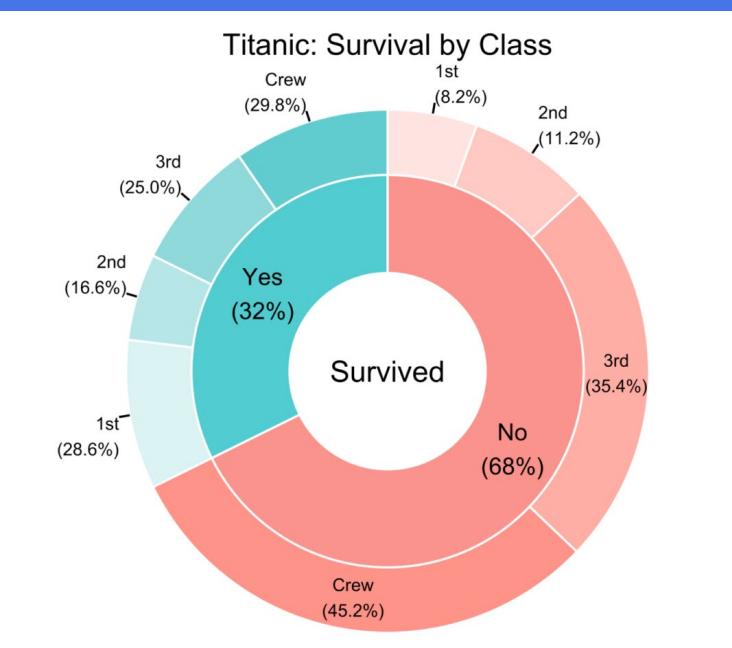




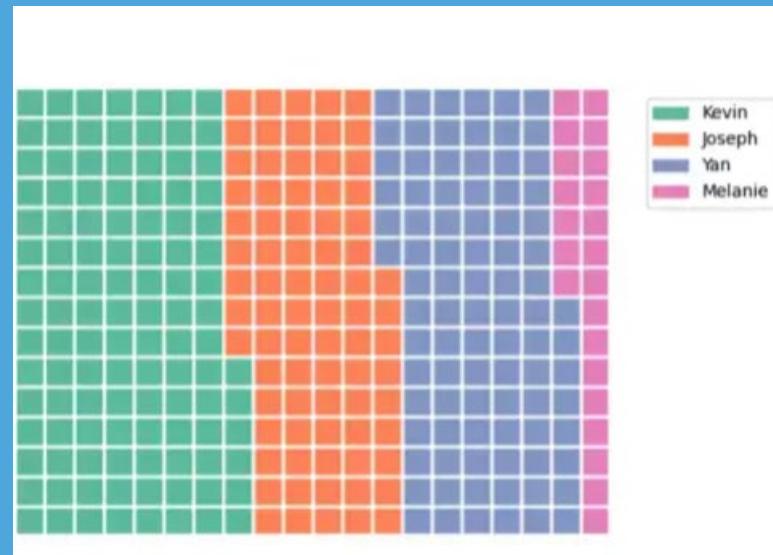
4. Proporciones - Composición



Treemap

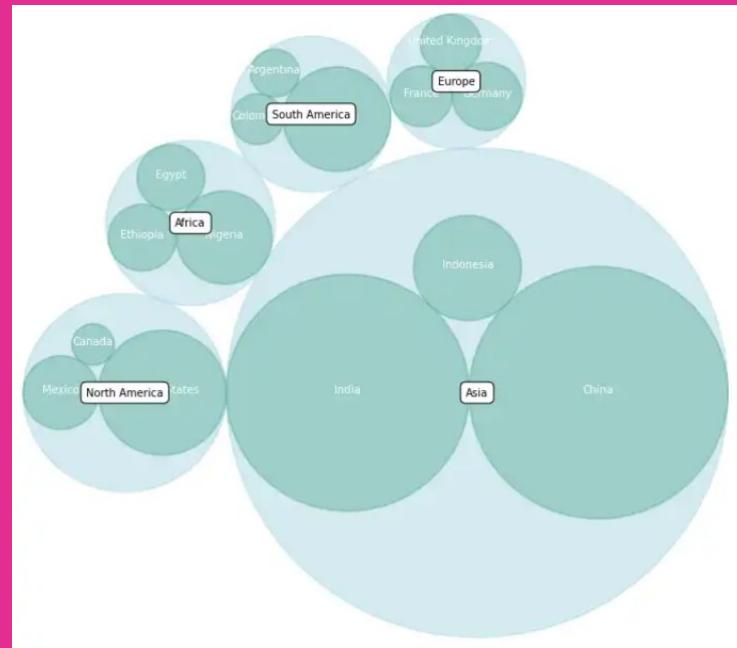


Donut Plot



Waffle Chart

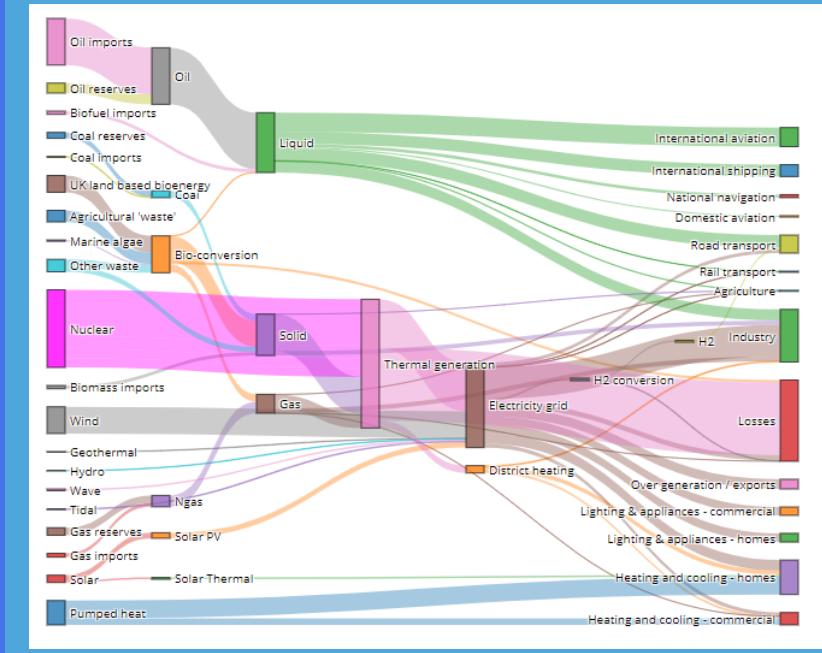
4. Proporciones - Composición



Circular Packing

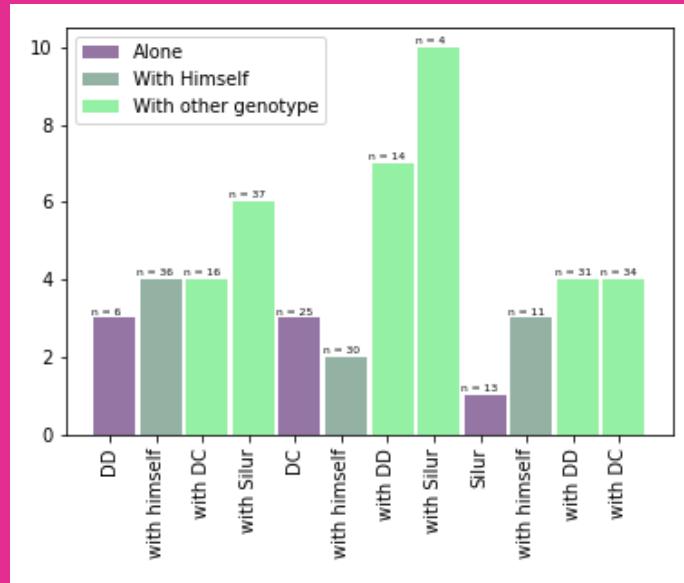


Funnel Chart

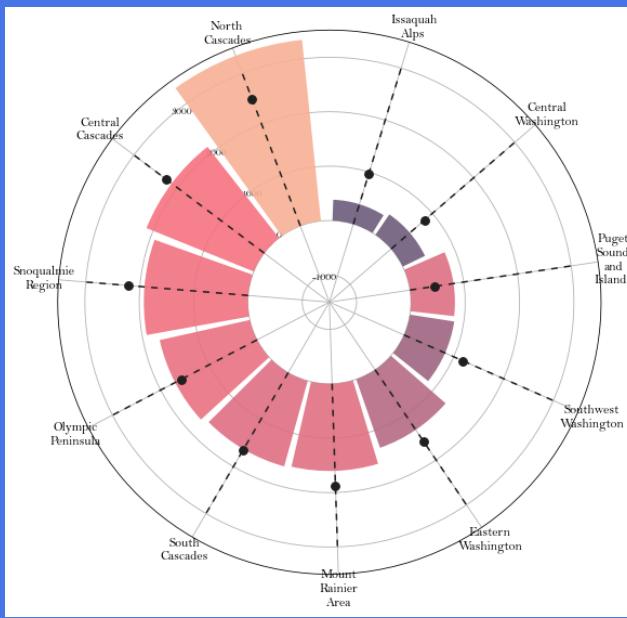


Sankey Chart

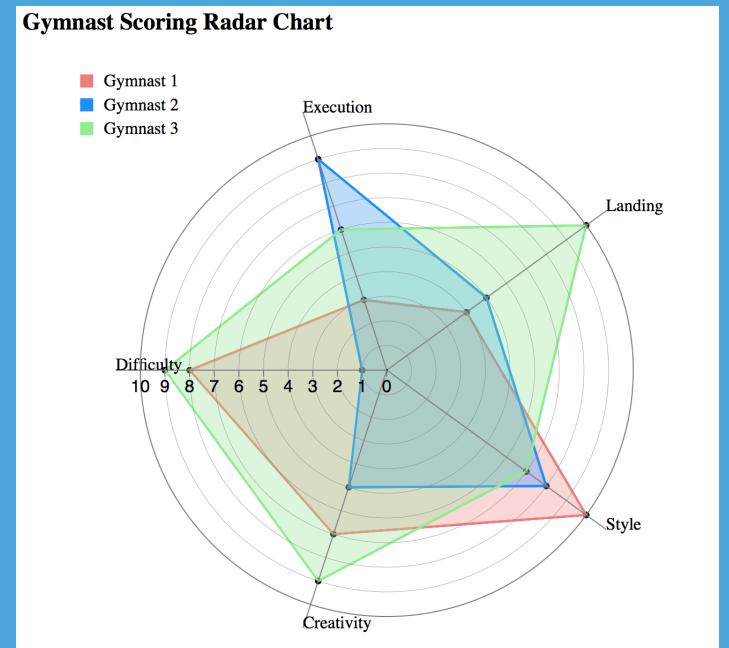
5. Ranking



Bar Plot



Circular Bar Plot



Waffle Chart

THE GLOBAL LIVEABILITY INDEX 2023

The Table visualizes a list of the Top 10 and Bottom 10 cities to Live in 2023.
It rates living conditions in 173 cities across these five categories: stability, health care, culture and environment, education and infrastructure.

City	Location	Rank	Index	Quality of Life			Environment		
				Stability	Healthcare	Education	Environment	Infrastructure	
Vienna		Austria	1	98.4	100.0	100.0	100.0	93.5	100.0
Copenhagen		Denmark	2	98.0	100.0	95.8	100.0	95.4	100.0
Melbourne		Australia	3	97.7	95.0	100.0	100.0	95.8	100.0
Sydney		Australia	4	97.4	95.0	100.0	100.0	94.4	100.0
Vancouver		Canada	5	97.3	95.0	100.0	100.0	97.2	96.4
Zurich		Switzerland	6	97.1	95.0	100.0	100.0	96.3	96.4
Calgary		Canada	7	96.8	100.0	100.0	100.0	87.3	100.0
Geneva		Switzerland	7	96.8	95.0	100.0	100.0	94.9	96.4
Toronto		Canada	9	96.5	100.0	100.0	100.0	94.4	89.3
Osaka		Japan	10	96.0	100.0	100.0	100.0	86.8	96.4
Auckland		New Zealand	10	96.0	95.0	95.8	100.0	97.9	92.9
Douala		Cameroon	164	46.4	60.0	29.2	41.7	51.2	42.9
Kyiv		Ukraine	165	44.0	40.0	41.7	75.0	54.2	23.2
Harare		Zimbabwe	166	43.8	40.0	29.2	66.7	56.7	35.7
Dhaka		Bangladesh	166	43.8	50.0	41.7	75.0	40.5	26.8
Port Moresby		Papua New Guinea	168	43.4	30.0	41.7	58.3	49.8	46.4
Karachi		Pakistan	169	42.5	20.0	50.0	75.0	38.7	51.8
Lagos		Nigeria	170	42.2	25.0	37.5	41.7	54.4	53.6
Algiers		Algeria	171	42.0	35.0	50.0	58.3	45.4	30.4
Tripoli		Libya	172	40.1	30.0	45.8	58.3	37.5	41.1
Damascus		Syria	173	30.7	20.0	29.2	33.3	40.5	32.1

Hazlo real

Entre mejor tus gráficas, mejor tu modelo

¿cómo es? ¿cómo sería?



**¿Quieres saber más?
¿Buscas códigos?**

<https://python-graph-gallery.com/>

<https://www.data-to-viz.com/>

<https://python-charts.com/es/>