



COMILLAS
UNIVERSIDAD PONTIFICIA

ICAI

comillas.edu

Universidad Pontificia Comillas ICAI

Analítica Social y de la Web *Social and Web Analytics*

1.1 Introducción a los datos sociales y de la web

Javier Ruiz de Ojeda

Curso 2023-24
Segundo semestre





COMILLAS
UNIVERSIDAD PONTIFICIA

ICAI

Presentaciones



Presentaciones

¿Quiénes sois vosotros?

- ¿Cómo te llamas? ¿De dónde vienes?
- ¿Por qué estudiar este Doble Grado? ¿Qué asignatura te ha gustado más?
- ¿Qué marcas o empresas te gustan más? ¿Cómo es su presencia online?
- ¿Qué tipo de puesto te gustaría desempeñar? ¿En qué tipo de empresa te gustaría hacerlo?
- ¿Tienes algún proyecto personal online? ¿Quieres contarnos?

Debate:

¿Qué esperáis de la asignatura?



Presentaciones

¿Y quién soy yo?





COMILLAS
UNIVERSIDAD PONTIFICIA

ICAI

Logística y datos de interés



Logística y datos de interés

¿Cómo nos vamos a organizar?

- Horario: lunes y miércoles de 16:00 a 18:00 (6 ECTS)
- Curso: 5º curso, segundo semestre
- Aula: A-109
- Contacto: Javier Ruiz de Ojeda (fjruizojeda@comillas.edu)

Logística y datos de interés

Índice de la asignatura

1. Proceso de Explotación de Datos
2. Obtención y Almacenamiento
3. Procesado de Datos
4. Análisis de sentimientos
5. Métricas de Experiencia de Usuario

Logística y datos de interés

Evaluación de la asignatura

- Exámenes: 50%
 - Inter: 15%
 - Final 35%
- Proyecto: 30%
- Ejercicios prácticos: 20%

Logística y datos de interés

Bibliografía de la asignatura

- **Siddhartha Chatterjee, Michal Krystyanczuk, Python Social Media Analytics, Packt Publishing 2017**
- Krishna Raj P.M., Ankith Mohan, Srinivasa K.G., Practical Social Network Analysis with Python, Springer 2018
- Matthew A. Russell, Mikhail Klassen, Mining the Social Web, 3rd Edition, O'Reilly 2019
- *(Ryan Mitchell, Web Scraping with Python, 2nd Edition, O'Reilly 2018)*
- *(Gabor Szabo, Gungor Polatkan, P. Oscar Boykin, Antonios Chalkiopoulos, Social Media Data Mining and Analytics, Wiley 2018)*



COMILLAS
UNIVERSIDAD PONTIFICIA

ICAI

¿De qué va esto?



¿De qué va esto?

Descripción “oficial”

“El objetivo de este curso introductorio a la Analítica Social y de la Web es el de proveer al estudiante con las herramientas teóricas y tecnológicas necesarias para la explotación y análisis de datos sociales generados en redes y en la web en general. El objetivo final es ganar experiencia práctica con casos de uso reales.”

¿De qué va esto?

Pues va de datos

INNOVATION

Data Is The New Oil -- And That's A Good Thing



Kiran Bhageshpur Forbes Councils Member

Forbes Technology Council COUNCIL POST | Membership (Fee-Based)

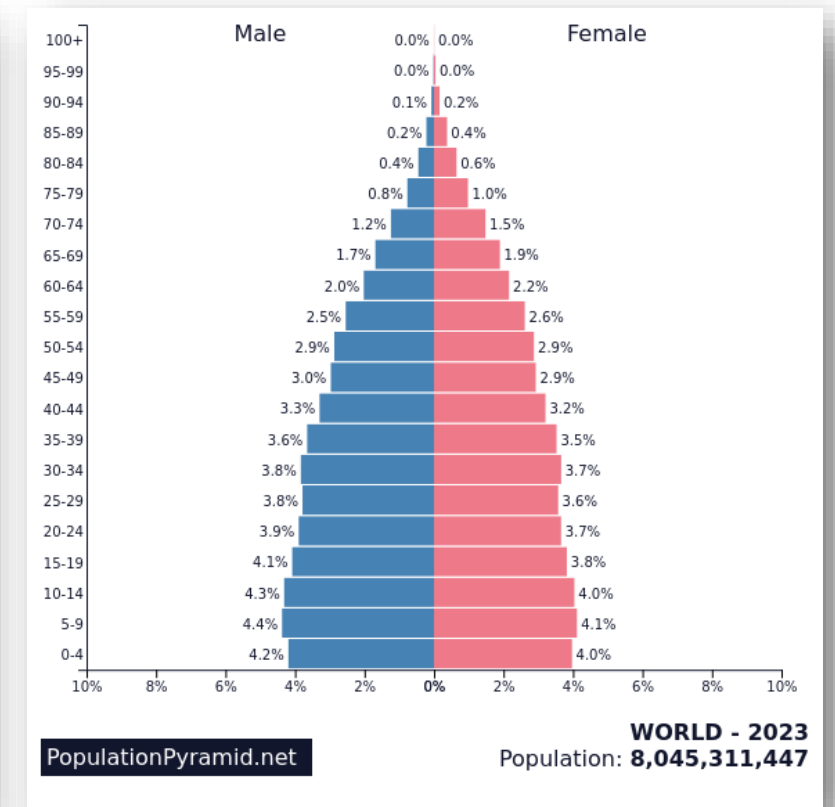
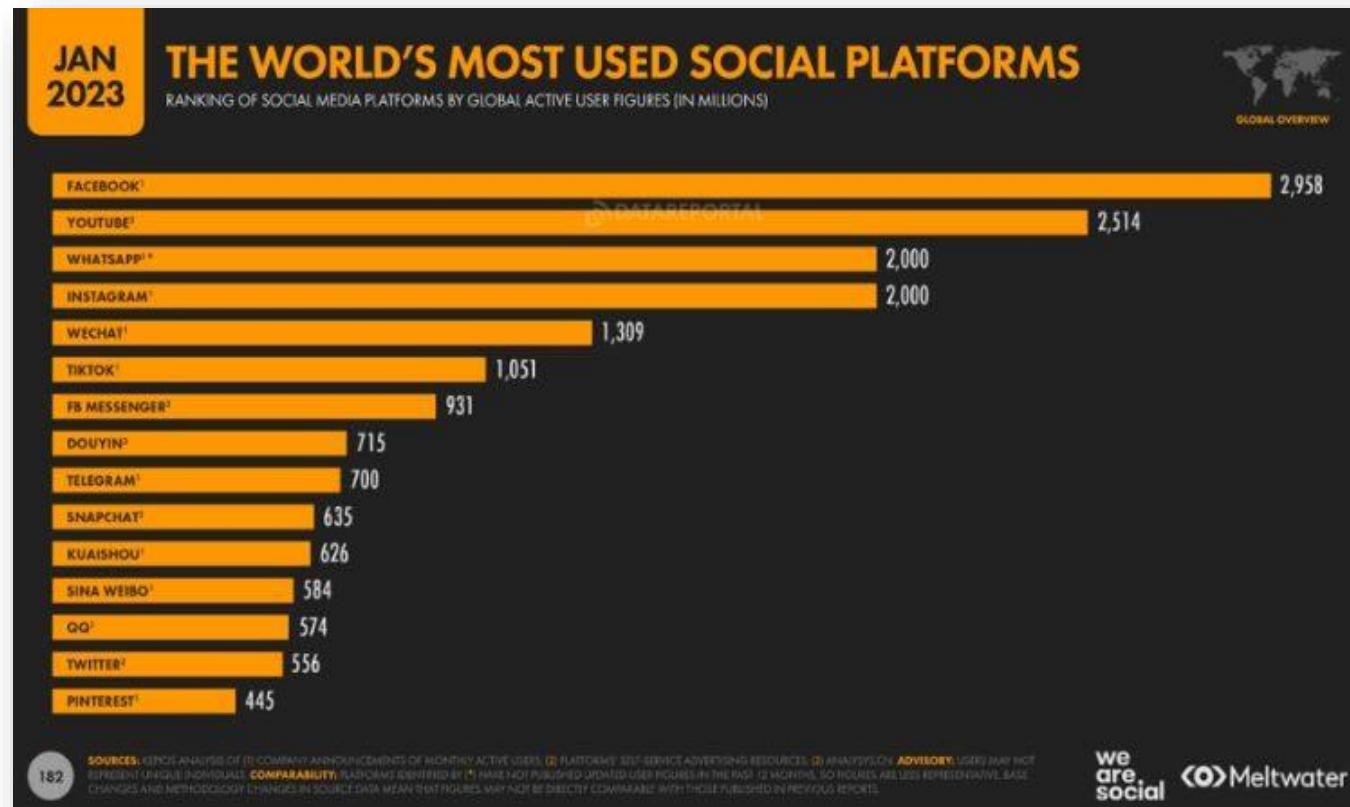
POST WRITTEN BY

Kiran Bhageshpur

CEO at Igneous. Reinventing [unstructured data management](#).

¿De qué va esto?

Y va de redes sociales

















¿De qué va esto?

¿Quién los usa?

Ranking de empresas por valor en Bolsa

Capitalizaciones a 28 de diciembre de 2023, en millones de dólares

Nº	EMPRESA	Sector	País	CAPITALIZACIÓN	Var. anual en %	Dif. puesto s/2022
1	Apple	 Tecnología	 EE UU	3.010.702	49,0	0 =
2	Microsoft	 Tecnología	 EE UU	2.789.179	56,5	1 ▲
3	Aramco	 Petrolera	 Arabia Saudí	2.132.800	14,3	-1 ▼
4	Alphabet	 Tecnología	 EE UU	1.761.588	58,9	0 =
5	Amazon	 Tecnología	 EE UU	1.585.034	82,6	0 =
6	Nvidia	 Tecnología	 EE UU	1.223.193	238,9	11 ▲
7	Meta (Facebook)	 Tecnología	 EE UU	920.833	197,8	18 ▲
8	Tesla	 Automoción	 EE UU	804.839	105,5	12 ▲
9	Berkshire Hathaway	 Finanzas	 EE UU	779.440	15,8	-3 ▼
10	Eli Lilly	 Salud	 EE UU	551.405	58,8	9 ▲



COMILLAS
UNIVERSIDAD PONTIFICIA

ICAI

Debate:

¿Qué es una red social y para qué sirve?

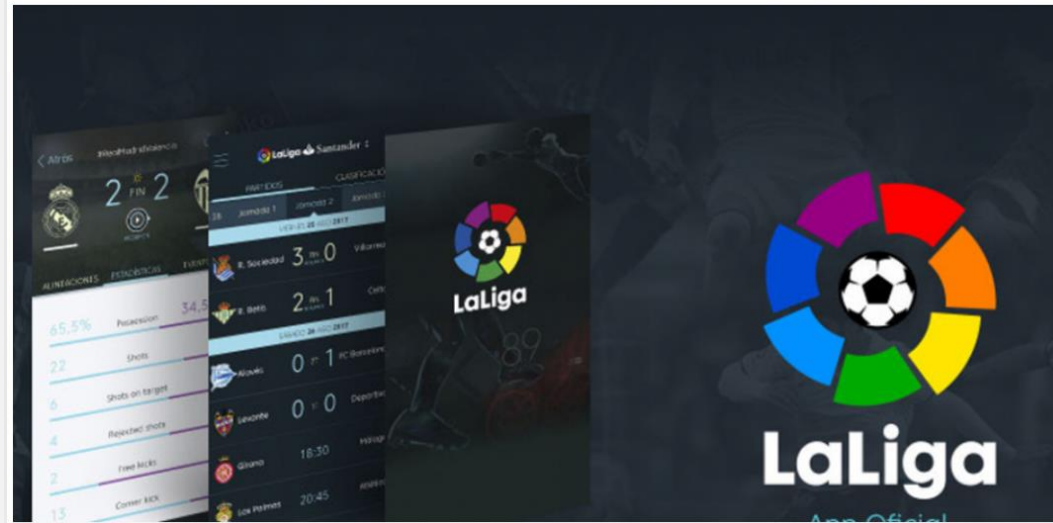


¿De qué va esto?

¿Y cómo los usa?

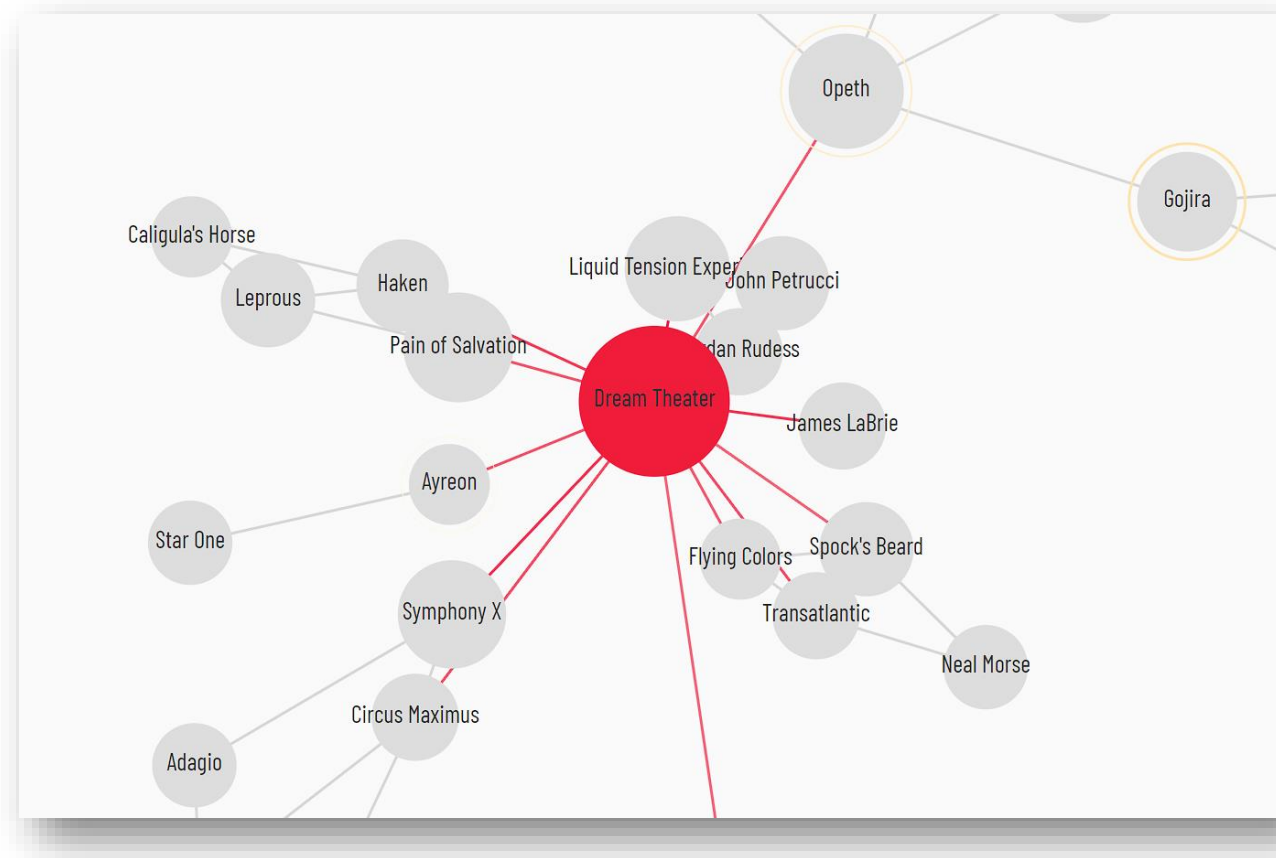
LaLiga, multada con 250.000 euros por falta de transparencia en su app

- A través del micrófono de la aplicación se recogían datos para perseguir la piratería
- LaLiga defiende que la tecnología utilizada ni graba, ni almacena, ni escucha conversaciones
- Aranda usaba el amaño de partidos para blanquear dinero del tráfico de drogas



¿De qué va esto?

También trata de representación y visualización



liveplasma.com

¿De qué va esto?

Trata de psicología

¿Puede la inteligencia artificial analizar los sentimientos y emociones de un tuit?

Twitter es un campo de análisis para la implementación de soluciones en seguridad, política, consumo y emociones.

Por: **The Conversation** 21 de abril del 2022 9:48 AM | Actualizado el 21 de abril del 2022 9:48 AM

Síguenos en Google News



Actualmente, es posible destilar la polaridad e intensidad mediante la identificación de rasgos léxicos, iconográficos y estructurales. | **Fuente:** Unsplash

COOSTO ESPAÑA



Inspiración ilimitada

Crea contenido que genere más notoriedad de marca, autoridad, fidelización y conversión.

¿De qué va esto?

Y en última instancia, trata de negocio

Así se hacen de oro las redes sociales con tus datos personales

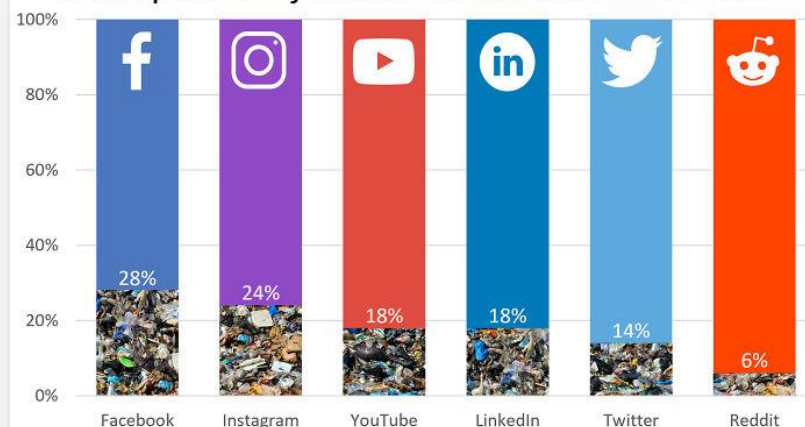
Este es el dineral que se embolsan por minuto las redes sociales con tus datos personales



Escrito por [Esther Lastra](#) 01 noviembre del 2022 a las 12:50

COMPARTIR [f](#) [t](#) [in](#) [✉](#)

¿Qué porcentaje de contenidos son anuncios?



*based on 100 home-page posts from each platforms' android app on my account. More information some surprises in comments.

Ingresos de las principales redes sociales por usuario

Tiempo medio diario		Por año	Por mes	Por semana	Por día	Por minuto	Por segundo
	33 Min.	201,10 €	16,75 €	3,85 €	0,55 €	1,00 €	0,017 €
	29 Min.	200,50 €	16,71 €	3,85 €	0,55 €	1,14 €	0,019 €
	32 Min.	65,89 €	5,49 €	1,26 €	0,18 €	0,34 €	0,006 €
	28 Min.	38,37 €	3,20 €	0,74 €	0,11 €	0,23 €	0,004 €
	31 Min.	53,63 €	4,47 €	1,03 €	0,15 €	0,28 €	0,005 €
	44 Min.	28,65 €	2,39 €	0,55 €	0,08 €	0,11 €	0,002 €

Datos de **heyData**

¿De qué va esto?

Entonces, ¿qué cualidades son deseables en los datos?

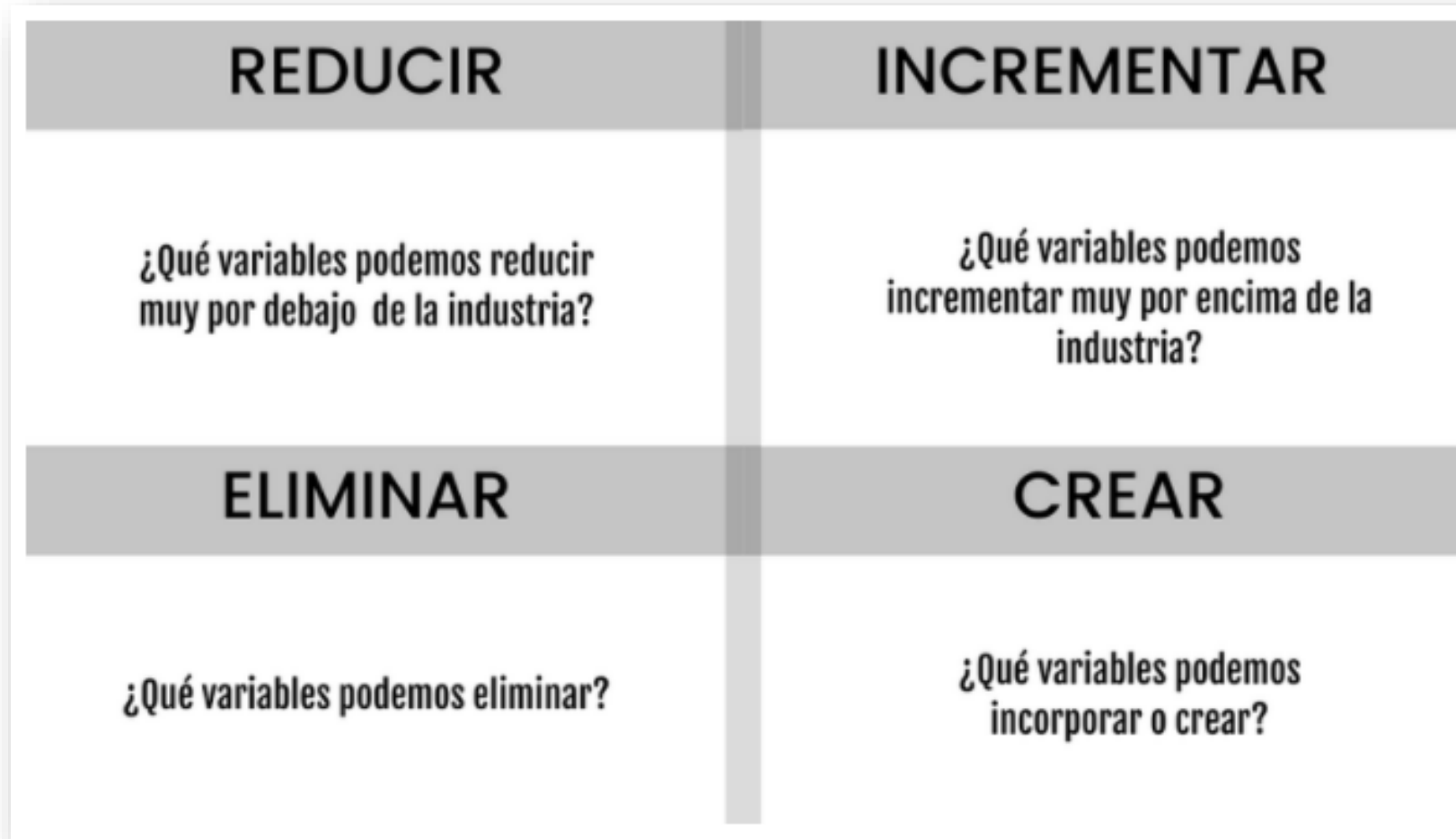
¿De qué va esto?

Desde un punto de vista de negocio

PREGUNTAS	MÉTRICAS	
¿Cuánto valor nos genera cada cliente en un mes o año?	ARPU Average Revenue per User	
¿Cuánto valor nos genera cada cliente?	$CLTV = \frac{\text{Ingresos generados por un cliente a lo largo de su vida}}{\text{Margen bruto (\%)}}$	
¿Cuánto (%) de clientes estamos perdiendo cada mes o año?	$CHURN\ RATE = \frac{\text{Clientes perdidos (en un periodo)}}{\text{Clientes totales (al principio)}}$	
¿Cuánto estamos invirtiendo para adquirir cada cliente?	$CAC = \frac{\text{Costes totales de marketing}}{\text{Nuevos clientes}}$	
¿Cuánto beneficio estamos obteniendo?	$CLTV - CAC$	$CLTV \& CAC\ Ratio = \frac{CLTV}{CAC}$
¿Cuánto tiempo vamos a tardar en recuperar lo invertido?	CAC Payback	

¿De qué va esto?

¿Cómo vamos a guiar nuestra toma de decisiones?



¿De qué va esto?

En resumen

Web Analytics hace referencia a la obtención, medición, análisis y reporte de los datos registrados en Internet con el fin de comprender optimizar el uso de una web en el ámbito del negocio.

De forma similar, **Social Media Analytics** es la obtención y análisis de datos de páginas web o aplicaciones de Redes Sociales para guiarnos en la toma de decisiones de negocio.

Por tanto, **la diferencia entre ambas no es más que la procedencia de los datos: Página web o Redes Sociales** de la empresa.

What Information can be gathered from?

Web Analytics

Four important key metrics can be analysed

1

Total Traffic

Total Traffic to your website gives insights about from where are you getting more traffic, that helps you to understand your target market.

2

Traffic Source

Traffic source is about how you are getting most visitors to your website.

@gectouch

3

Bounce Rate

Bounce rate is the percentage of visitors to a particular website who are leaving the site after viewing only one page without navigating other pages on the website.

4

Conversion Rate

A conversion rate is the percentage of visitors who have taken some action in your website or complete the desired goal.

Social Media Analytics

Social media analytics is to discovery

1

Sentiments

It is the process of computationally identifying and categorizing opinions expressed in a piece of text.

2

Opinions

Opinion about the product or services and review will be shared by the consumers in social media networks.

3

Emotions

Mostly consumer will share their emotion about the product only if they are extremely satisfied or dissatisfied.

4

Topics

Analysing the topic in the social media will give you good insight about market trends and expectation.

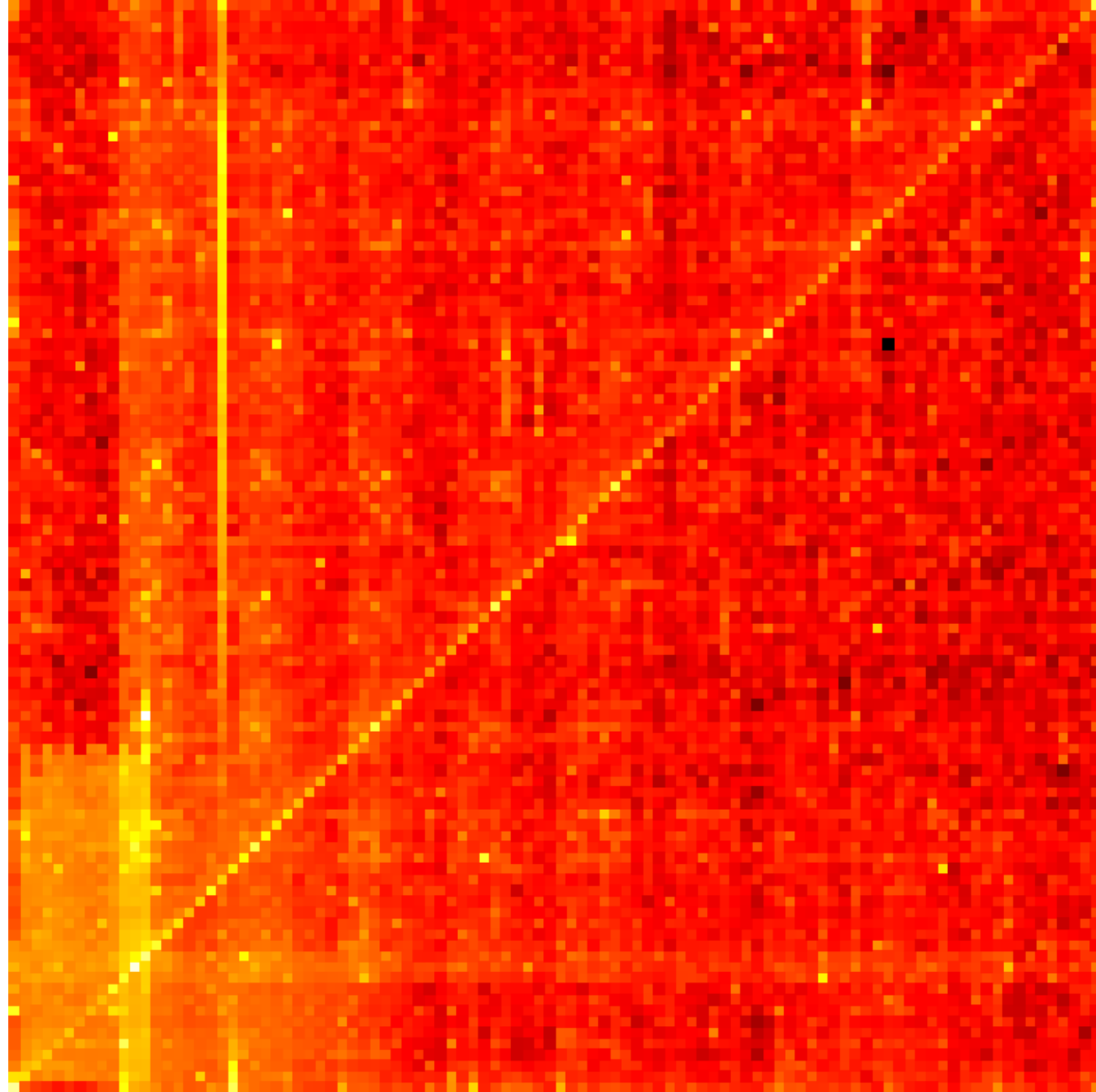


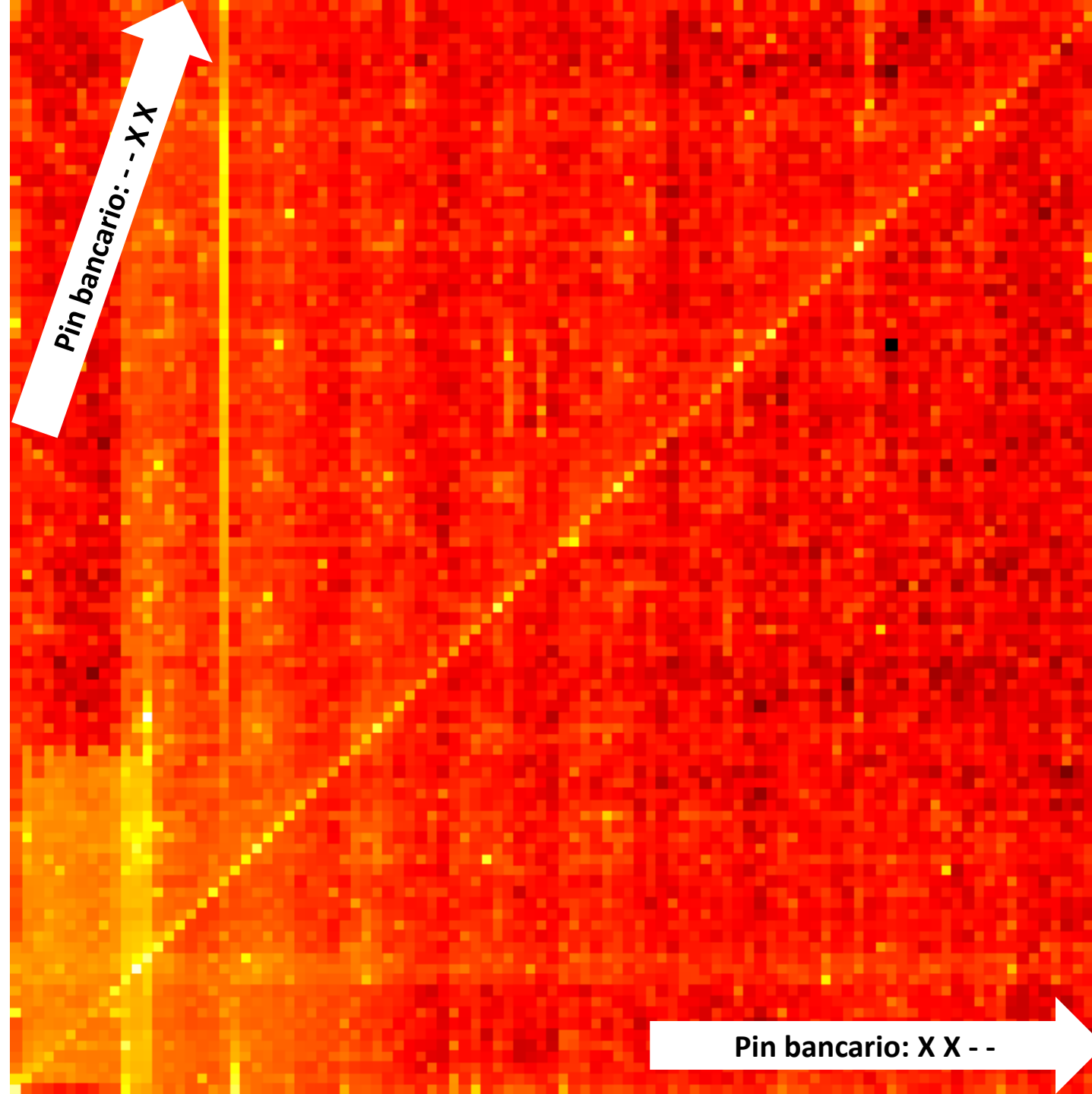
COMILLAS
UNIVERSIDAD PONTIFICIA

ICAI

Prueba de nivel... ⇨⇨

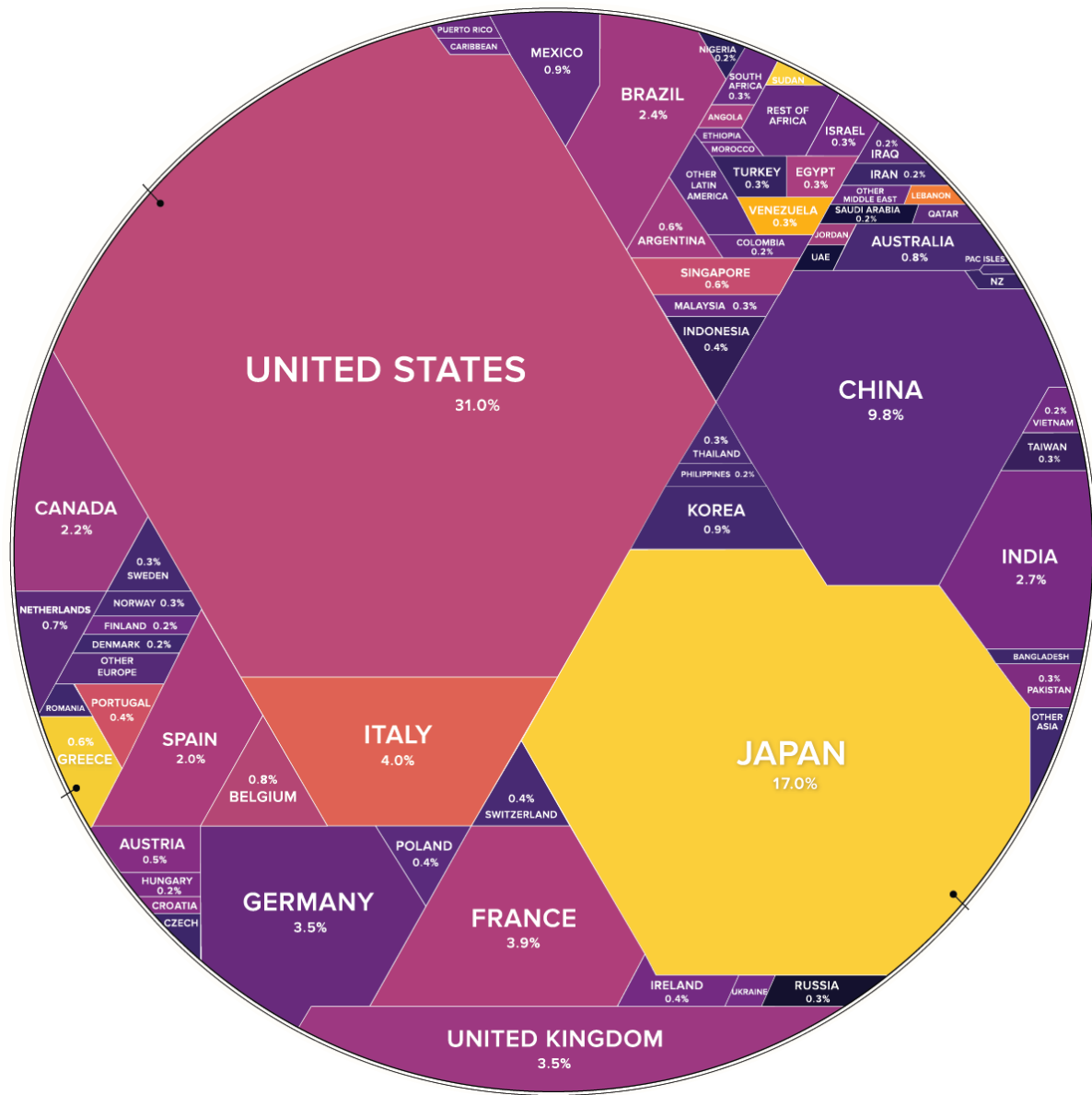




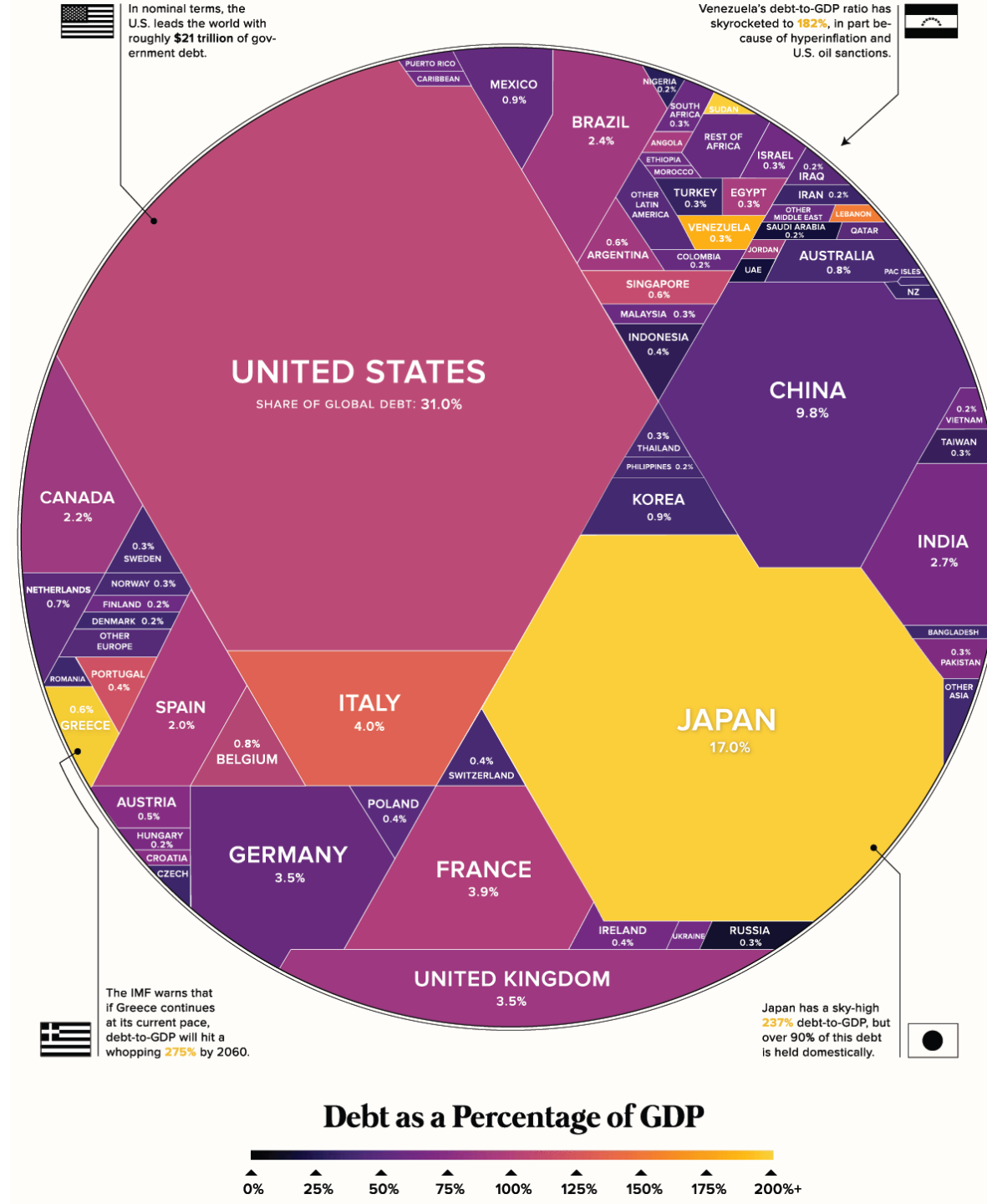


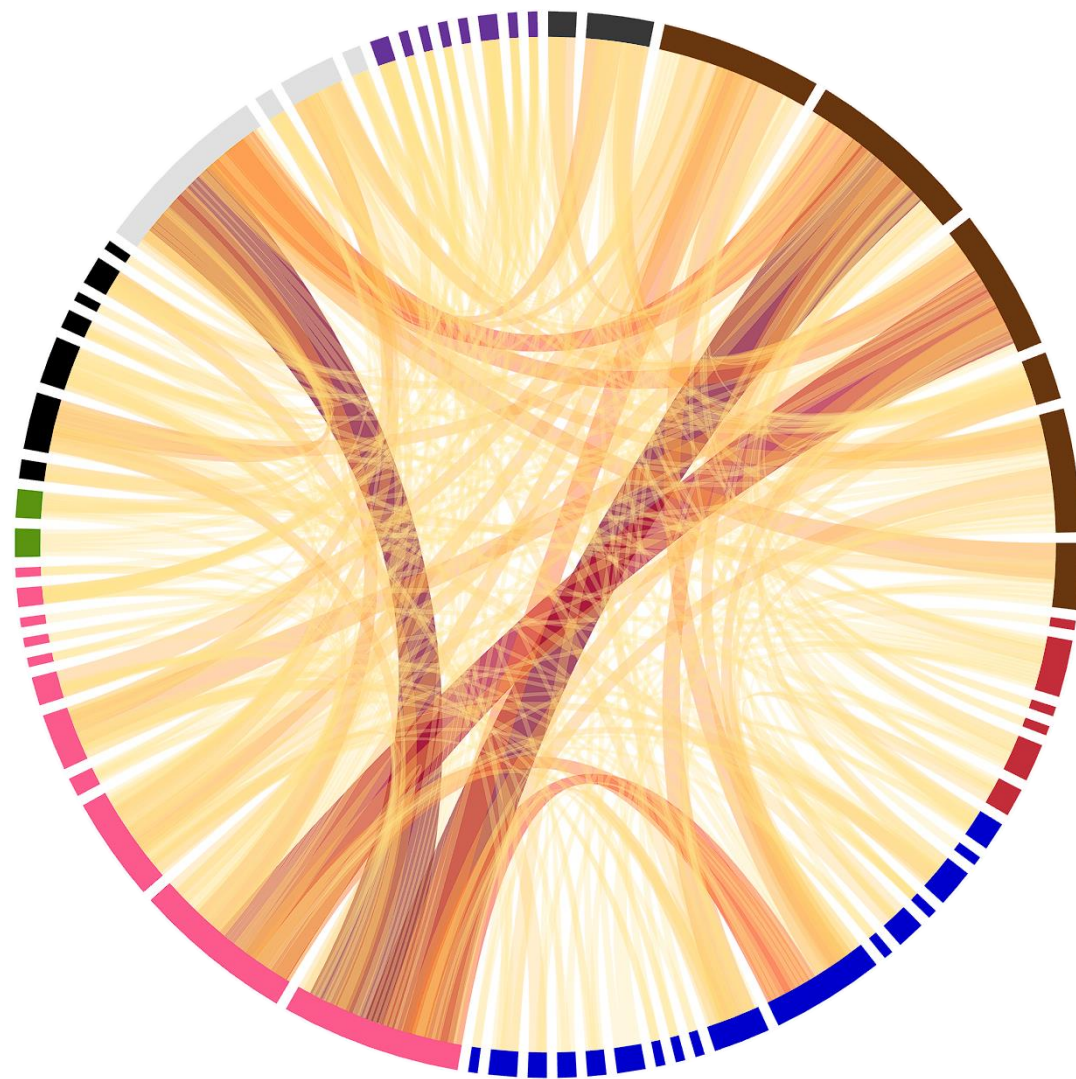
Pin bancario: - - X X

Pin bancario: X X - -

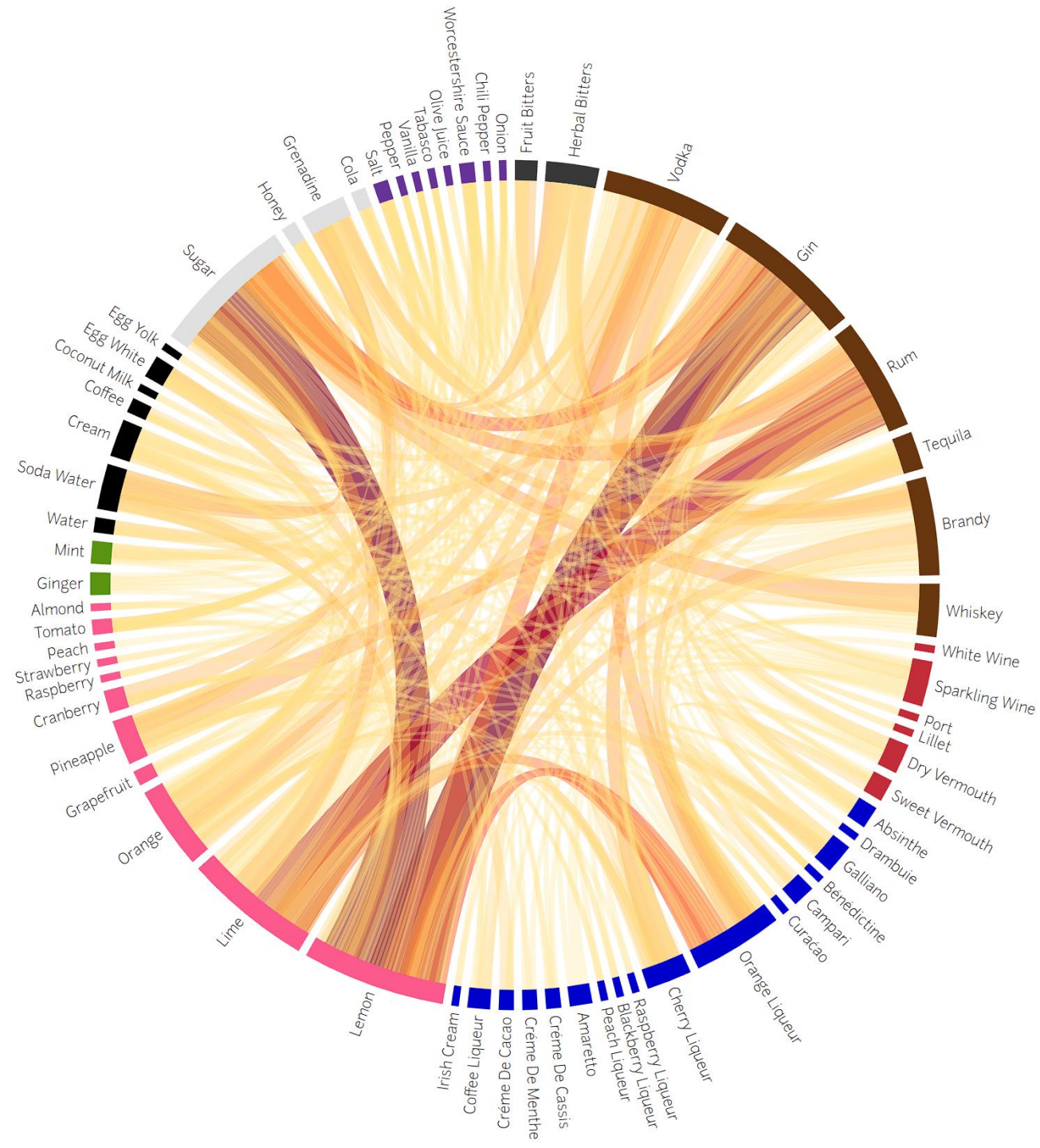


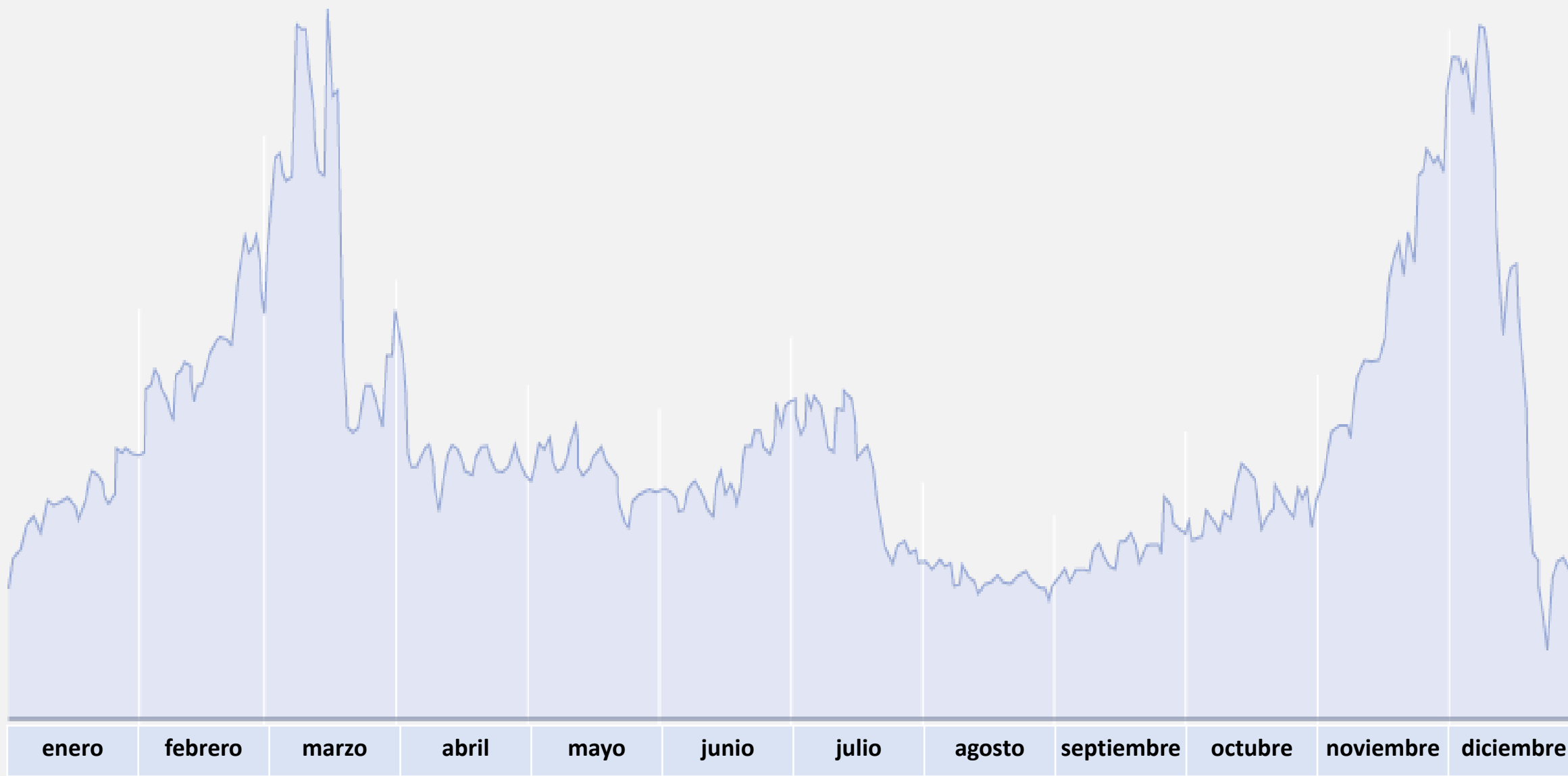
Nivel de deuda como parte del global de cada país



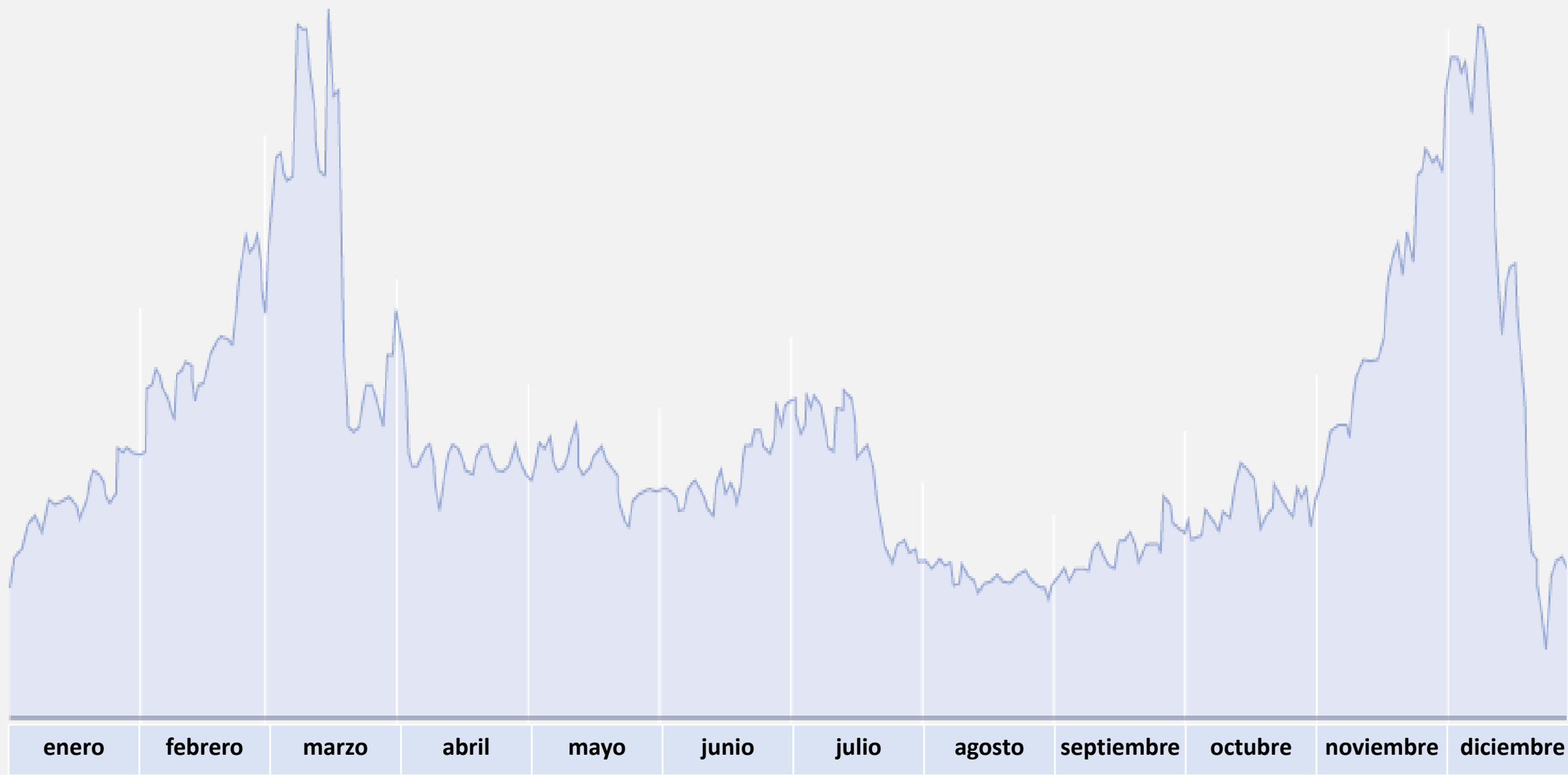


Combinatoria estadística entre ingredientes en la coctelería mundial





Rupturas de parejas según los estados de Facebook





COMILLAS
UNIVERSIDAD PONTIFICIA

ICAI

Business Case:



Cambridge
Analytica

Analytica

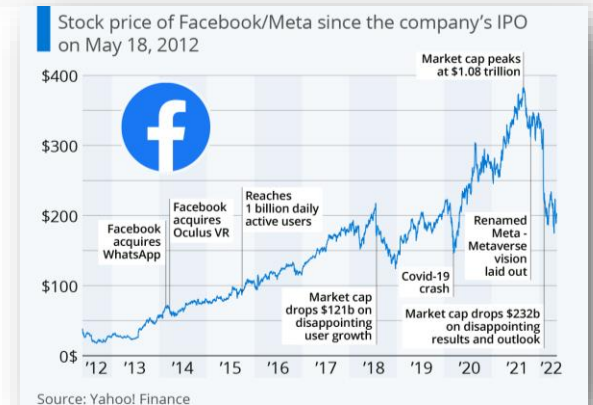
Cambridge



Business Case: *Cambridge Analytica*

Cronología de un escándalo

- 2004: Nace Facebook como red universitaria
- 2008: Comienza su liderazgo mundial en RRSS
- 2012: Comienza a cotizar en bolsa. Valor: 100.000 millones
- 2012: Alcanza los 1.000 millones de usuarios
- 2014: Compra de Instagram por 1.000 millones
- 2014: Compra de Whatsapp por 21.800 millones
- 2014: Cambridge Analytica lanza una aplicación para obtener datos de usuarios a través de Facebook



Business Case: *Cambridge Analytica*

Cronología de un escándalo



- 2015: Facebook descubre que Cambridge Analytica ha obtenido datos no solo de los usuarios que han usado la app, sino de todos los contactos de cada participante de manera exponencial
- 2016: Donald Trump gana las elecciones a la presidencia de EEUU. Cambridge Analytica trabaja como consultora para su candidatura
- 2017: Se sospecha de la utilización de Fake News a través de las RRSS para manipular las elecciones y el resultado del Brexit
- 2018: 17 de Marzo. Se publica en prensa que Cambridge Analytica llegó a obtener hasta 50 millones de datos de usuarios para utilizar en la campaña electoral y en el brexit

Business Case: *Cambridge Analytica*

Cronología de un escándalo

- 2018: 18 de Marzo. Facebook suspende la cuenta de Cambridge Analytica
- 2018: 20 de Marzo. Se solicita desde el Parlamento Británico la comparecencia de Mark Zuckerberg y Cambridge Analytica
- 2018: 21 de Marzo. Las pérdidas en bolsa de Facebook alcanzan los 50.000 millones
- 2018: 21 de Marzo. Mark Zuckerberg publica a través de su perfil de Facebook un extenso comunicado explicando la situación
- 2018: 22 de Marzo. Se presentan diferentes demandas particulares ante la Corte Federal de San Francisco en EEUU por la pérdida del valor de las acciones

Business Case: *Cambridge Analytica*

Cronología de un escándalo

- 2018: 22 de Marzo. Se viraliza el Hashtag en Twitter #DeleteFacebook
- 2018: 22 de Marzo. El Congreso de EEUU también llama a declarar a Mark Zuckerberg
- 2018: 22 de Marzo. Facebook anuncia la contratación de un equipo forense para determinar si Cambridge Analytica todavía posee los datos
- 2018: 3 de Abril. Las pérdidas alcanzan los 100.000 millones
- 2018: 4 de Abril. Facebook confirma ante la prensa que los datos robados ascienden a 87 millones
- 2018: 4 de Abril. Facebook confirma También que escanean y analizan los enlaces y las fotos de los usuarios de Messenger

Business Case: *Cambridge Analytica*

Cronología de un escándalo

- 2018: 11 de Abril. Mark Zuckerberg testifica bajo juramento ante el Comité de Comercio y Energía de la Cámara de Representantes de EEUU



N PELÍCULA

EL GRAN HACKEO



Reproducir





COMILLAS
UNIVERSIDAD PONTIFICIA

ICAI

Debate:

*¿Qué datos ofrece cada red social
y para qué nos pueden servir?*





COMILLAS
UNIVERSIDAD PONTIFICIA

ICAI

Ejercicio:

*¿Qué datos ofrece cada red social
y para qué nos pueden servir?*





COMILLAS
UNIVERSIDAD PONTIFICIA

ICAI

¡Muchas gracias!

