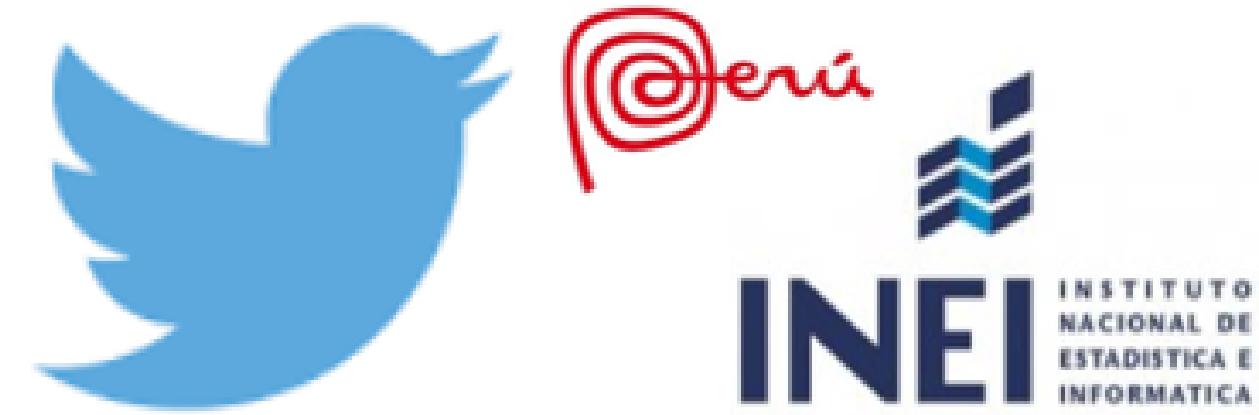


Análisis de ofertas de empleo mediante datos obtenidos de Twitter e información del INEI



INTEGRANTES:

- MIJAIL DAVIS HUAMAN ROMERO
- EDUARDO HUARCAYA QUISPE
- VICTOR ANDRES CORDOVA BERNUY

Caso de Uso

El presente trabajo tiene por finalidad analizar la relación existente entre las ofertas de empleo publicadas en la red social Twitter versus el tipo de carreras universitarias y técnicas estudiadas por los limeños.

El estudio tendrá como base 04 pilares:

- Procesamiento de información a través de tecnologías Big Data
- Procesamiento de información a través de algoritmos de aprendizaje autónomo “Machine Learning”
- Recolección masiva de información de la web a través del uso de algoritmos de “Web Scraping”
- Visualización de datos a través de herramientas como Power BI y Excel.

Necesidad de Tecnologías Big Data



El presente trabajo sustentara su análisis con información proveniente de dos grandes fuentes de datos.

- ❖ **Twitter**, considerada una de las mas importantes redes sociales de la actualidad con 330 millones usuarios activos y 500 millones de Tweets por día (*).
- ❖ **INEI**, Instituto Nacional de Estadísticas e informáticas, organismo encargado de los censos de población, empresas, viviendas, etc., en Perú.

Dada la magnitud de información que es posible recolectar se nos hace viable la aplicación de Tecnologías Big Data.

(*) Fuente: <https://www.flimper.com/blog/es/estadisticas-globales-de-twitter-2018->

Análisis de los Datos (3V's)

- ❑ **Velocidad:** Al tener como fuente de información datos proveniente de Twitter, estos se generan constantemente a través de Tweets publicados.
- ❑ **Variedad:** Se procesara información proveniente del INEI (data estructurada) e información proveniente de Twitter (data no estructurada).
- ❑ **Volumen:** INEI y Twitter, gestionan grandes volúmenes de información.

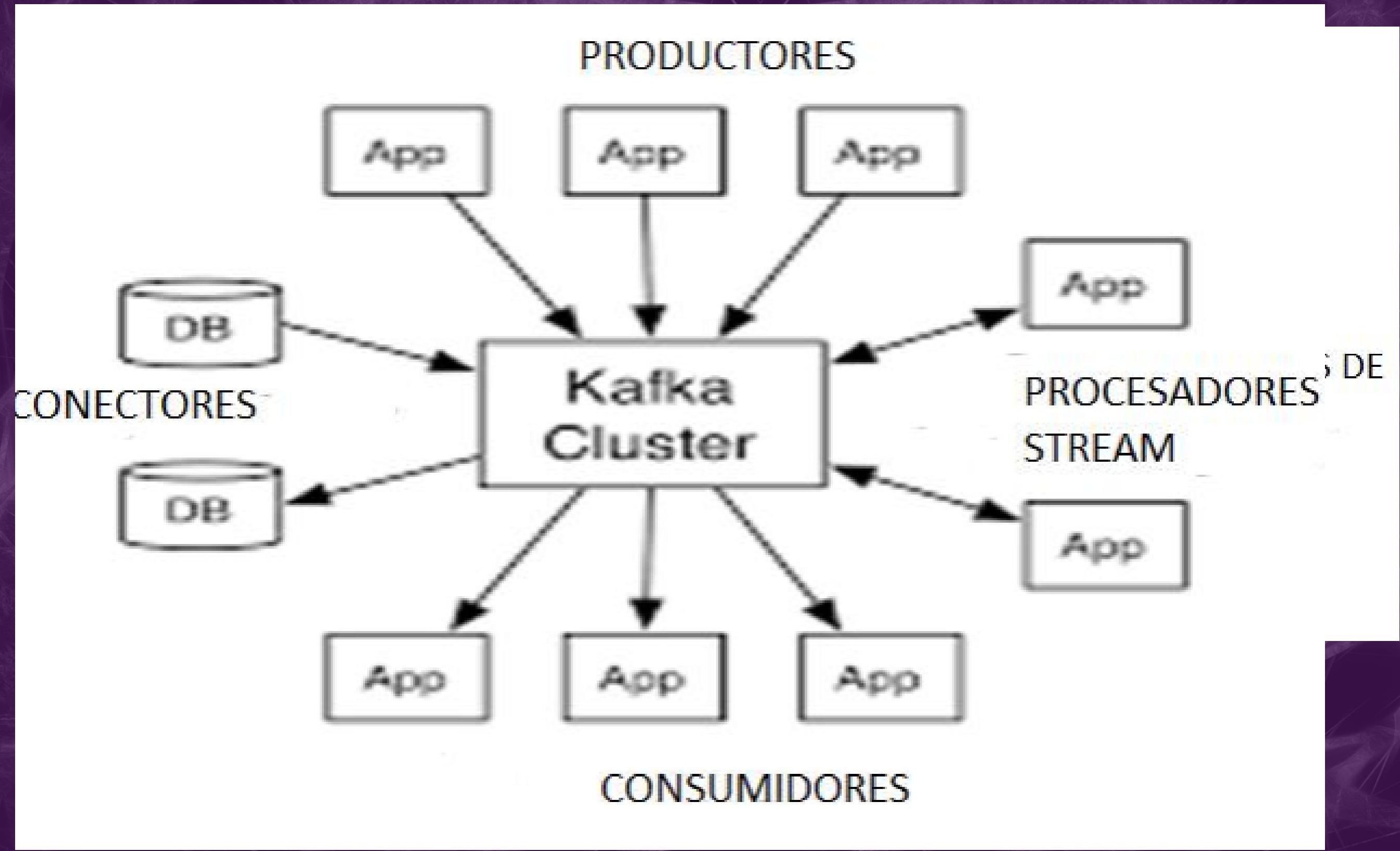
ARQUITECTURA STREAMING CON KAFKA

A nivel de diagrama de Caja Nagra



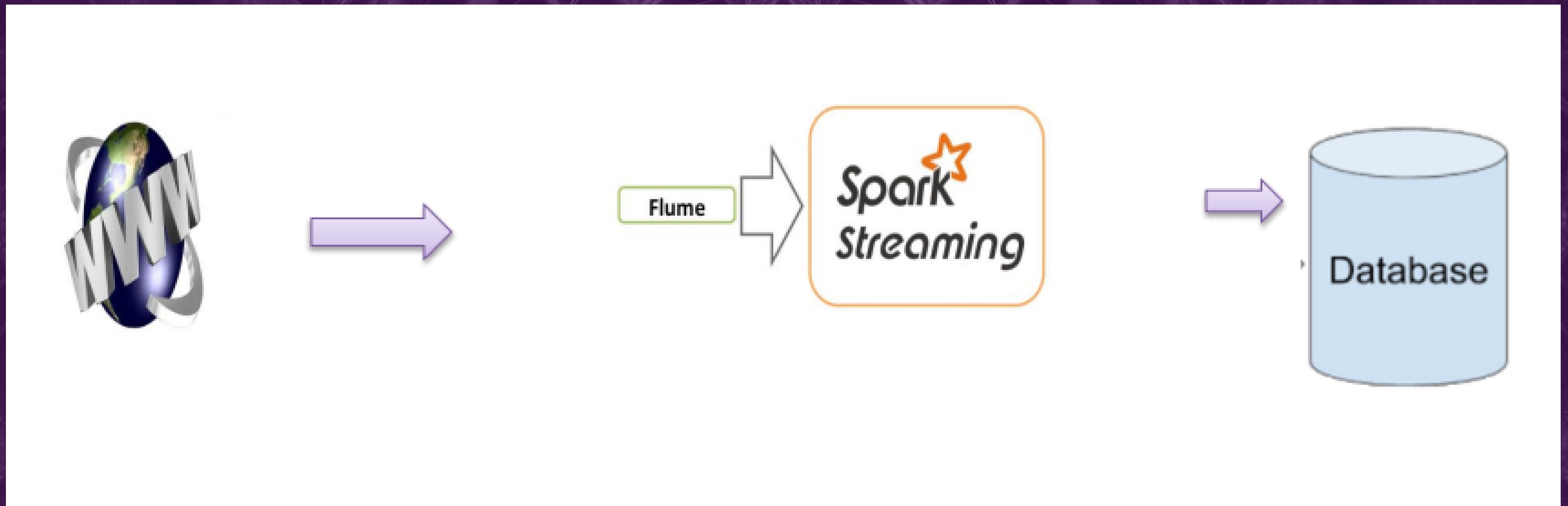
ARQUITECTURA STREAMING CON KAFKA

A nivel de caja blanca



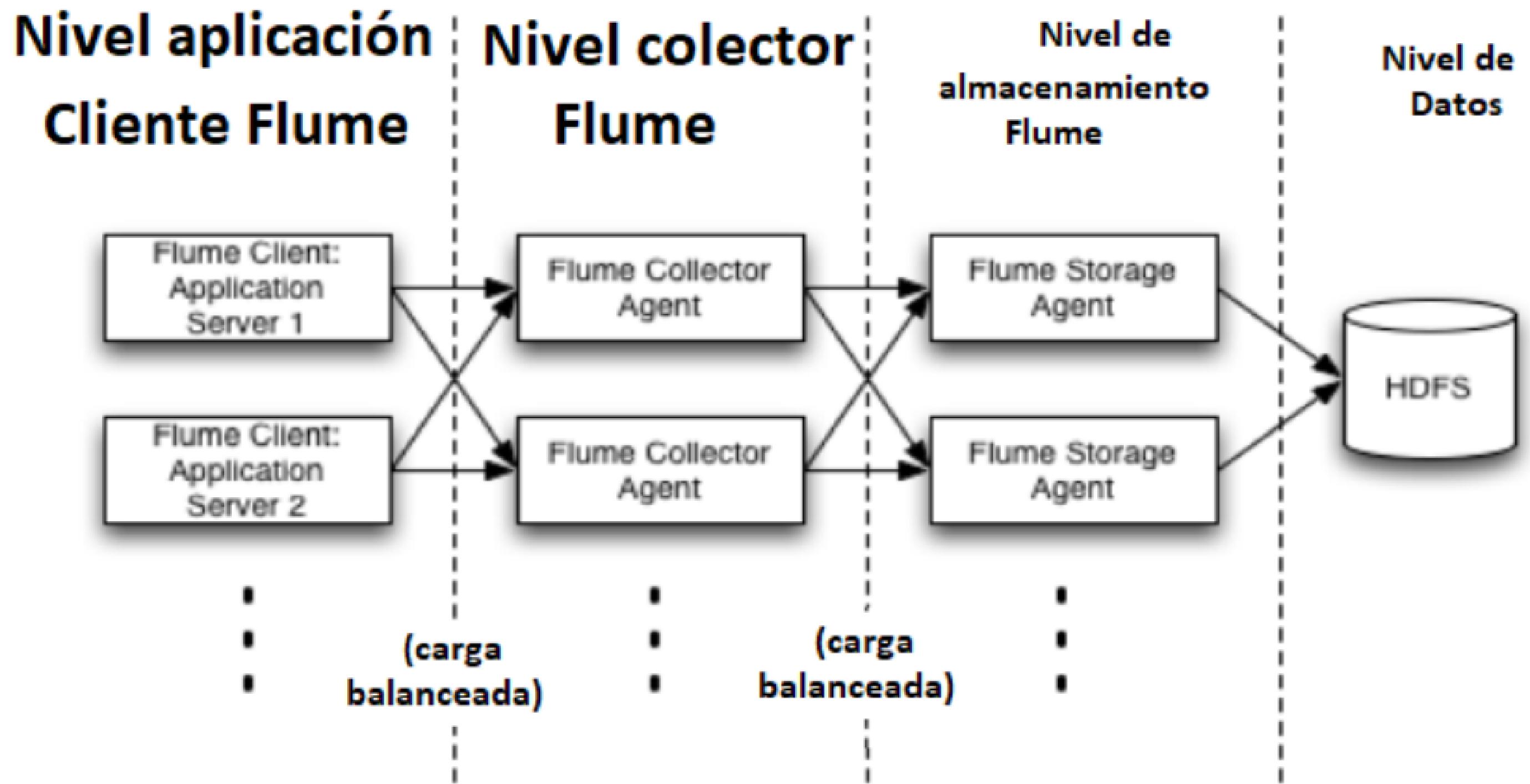
Arquitectura Streaming con Apache Flume

A nivel de caja negra



A nivel de diagrama de caja blanca

Una topología de bloques Flume



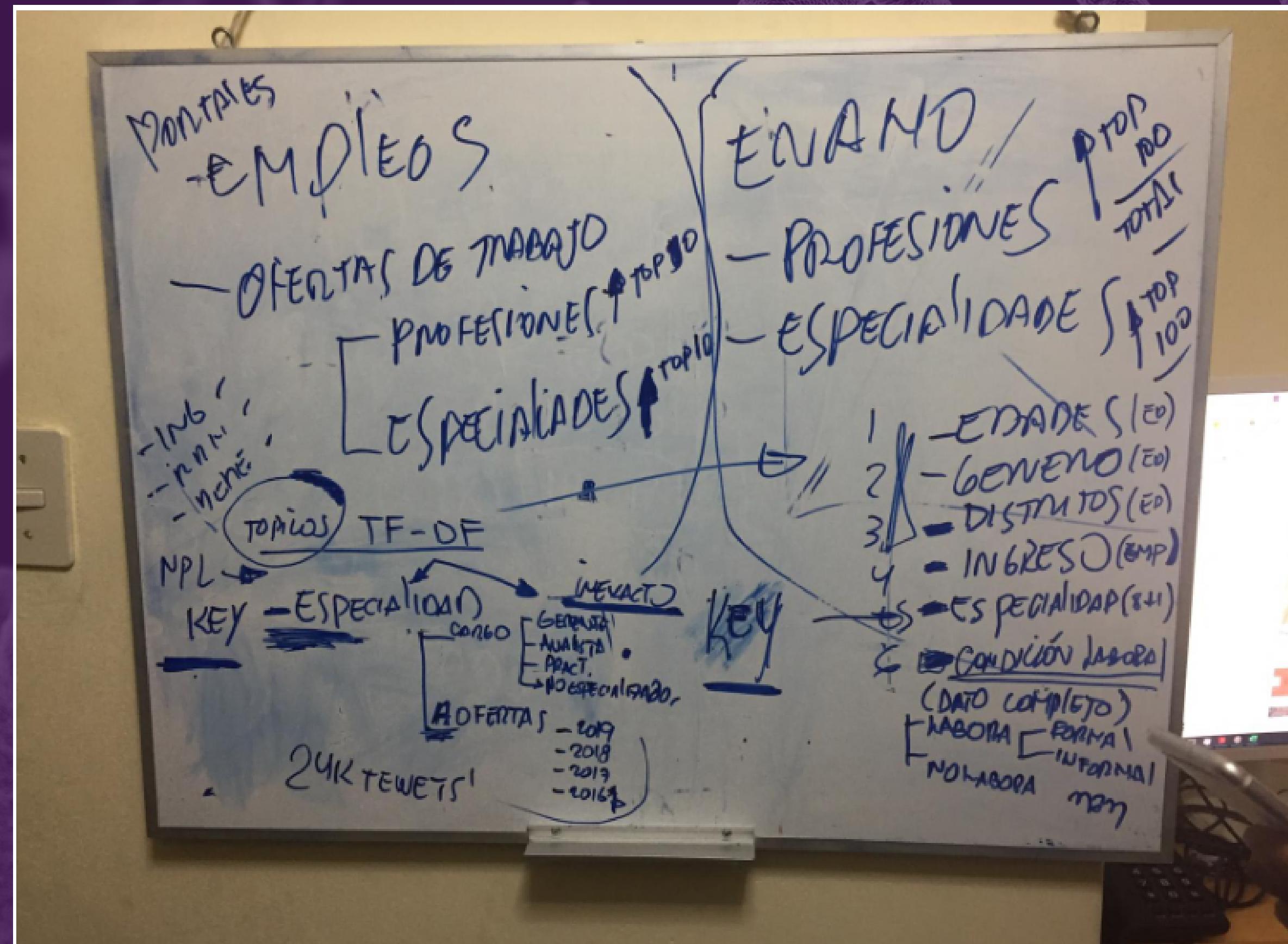


DIAGRAMA DE POWER BI

Sales Home View Modeling Help

Get Data Refresh Sources Enter Data Refresh New Page New Visual Ask A Question Insert Tool Box Shapes Power Marketplace Power File Switch Theme Manage Relationships New Measure New Column New Quick Measure Calculations Publish

Clipboard External data Insert Custom visuals Themes Relationships Share

Eduardo Huarcaya

Count of Twitter and Count of Twitter by year and Portal

Portal: applusemployees, Labourunpe, MargoverGroupIT, ovialbusiness, peru_trabajo, perusemployees, Count of Twitter

Portal: applusemployees, Labourunpe, MargoverGroupIT, ovialbusiness, peru_trabajo, perusemployees, Count of Twitter

Count of Twitter and Count of Twitter by month and Portal

Portal: applusemployees, Labourunpe, MargoverGroupIT, ovialbusiness, peru_trabajo, perusemployees, Count of Twitter

Visualizations > Fields >

Search

Values

Add data fields here

Filters

Page level filters

Add data fields here

Report level filters

Add data fields here

Drillthrough

Cross-report

Off —

Keep all filters

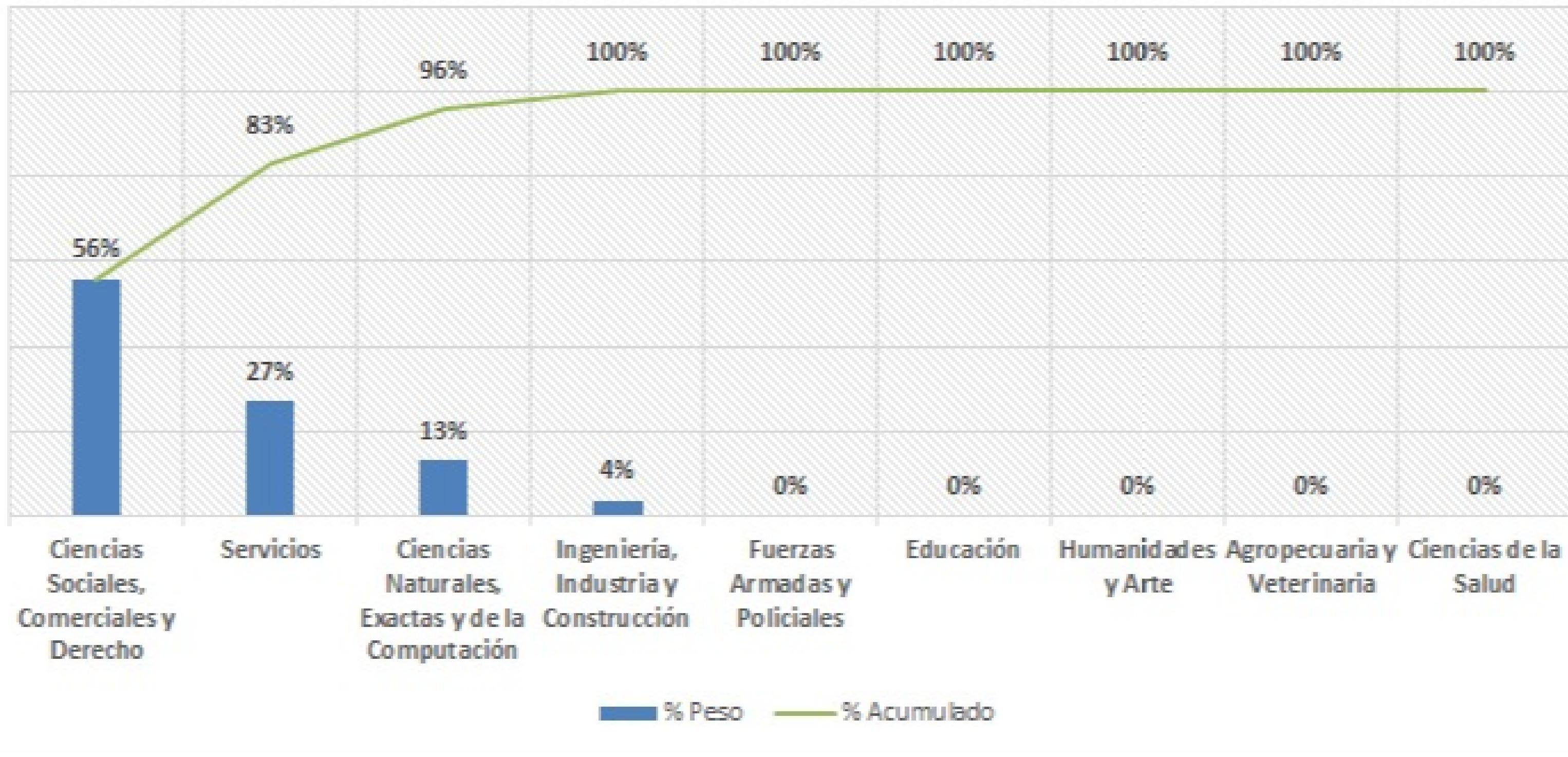
On —

Add drillthrough fields here

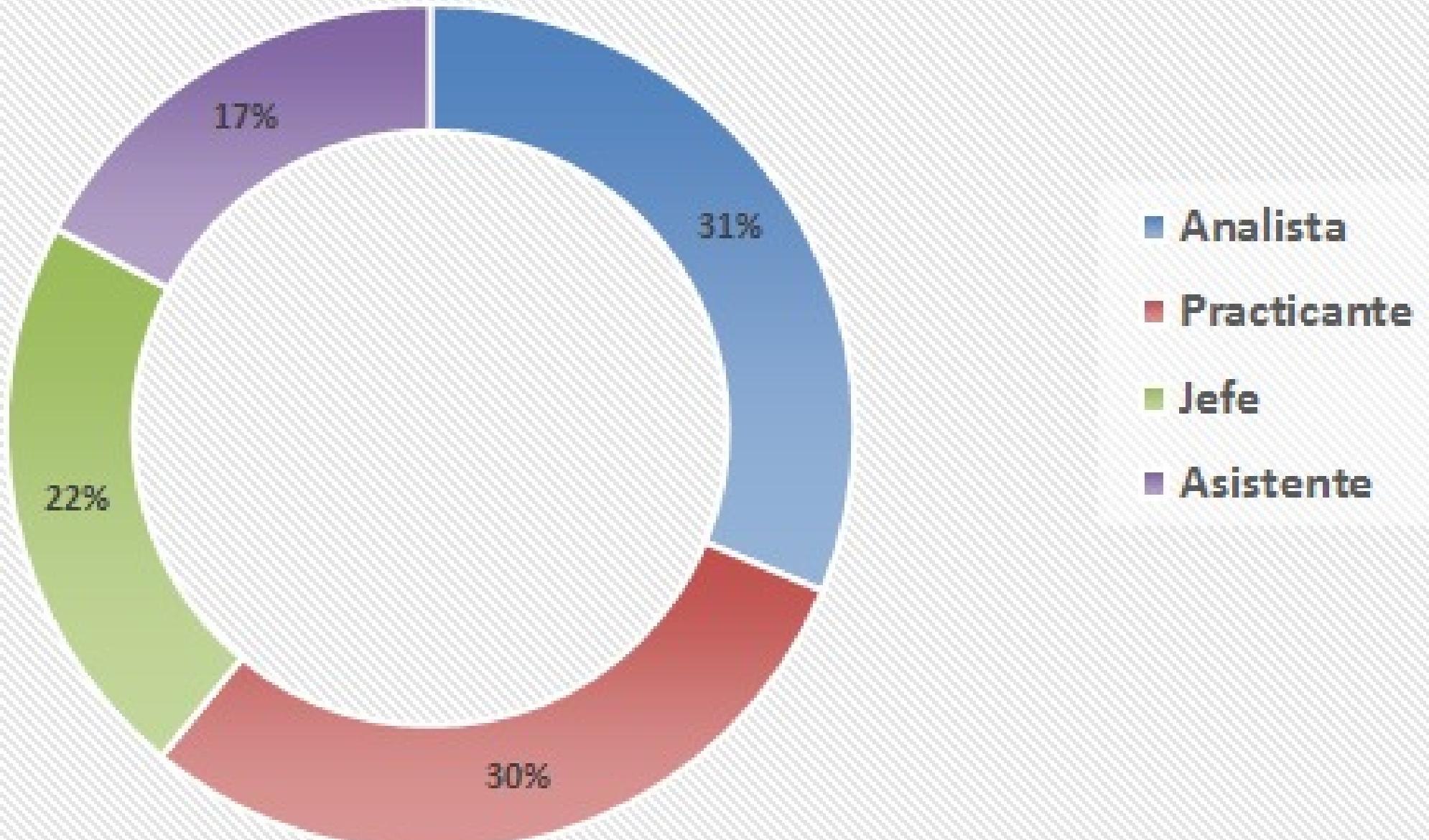
Page 1 +

PAGE 1 OF 1

Especialidades más demandadas en la Web



Cargos más demandadas en Twitter



Regla Barra de fórmulas
Líneas de cuadrícula Títulos

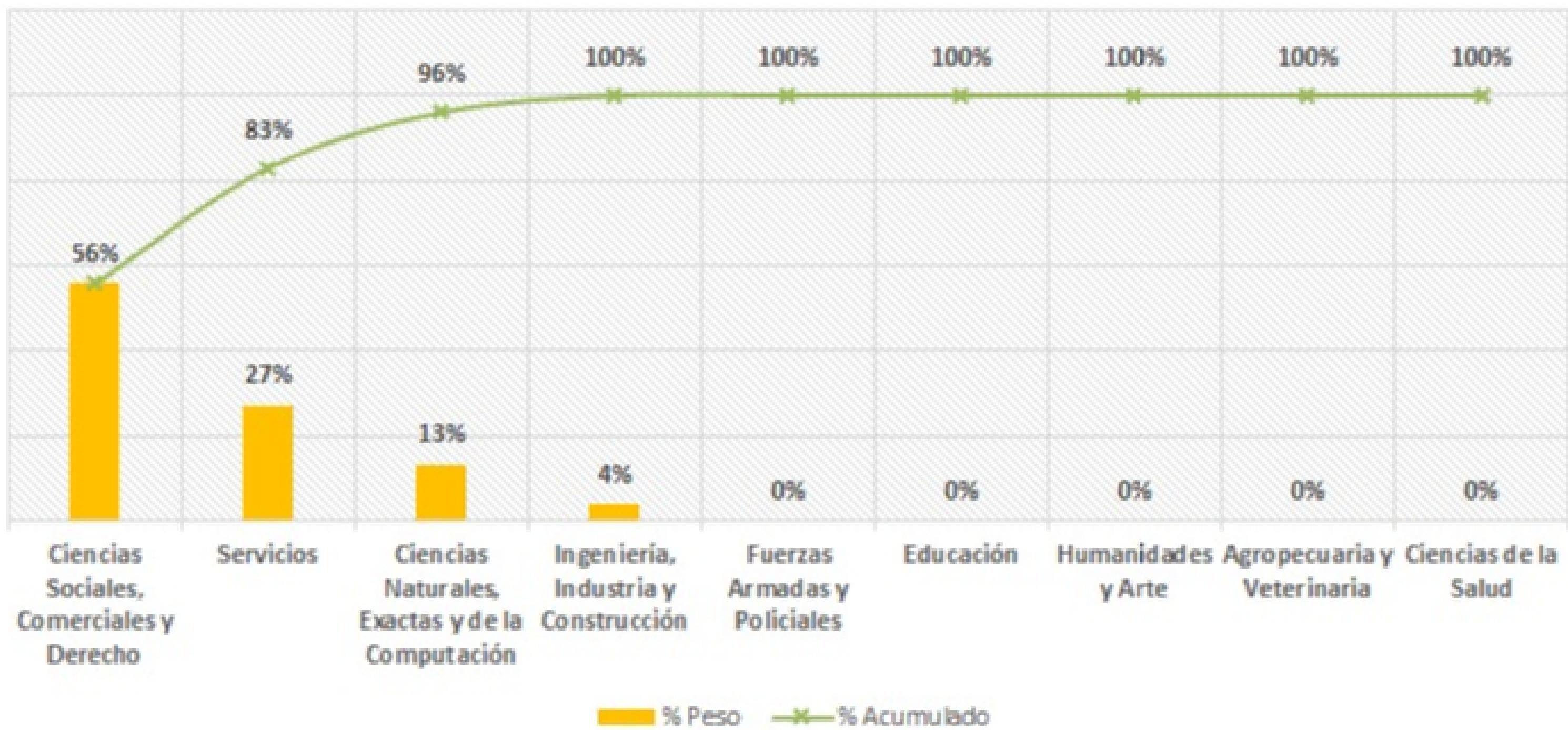
Zoom 100% Ampliar selección Nueva Organizar Inmovilizar

Mostrar Zoom

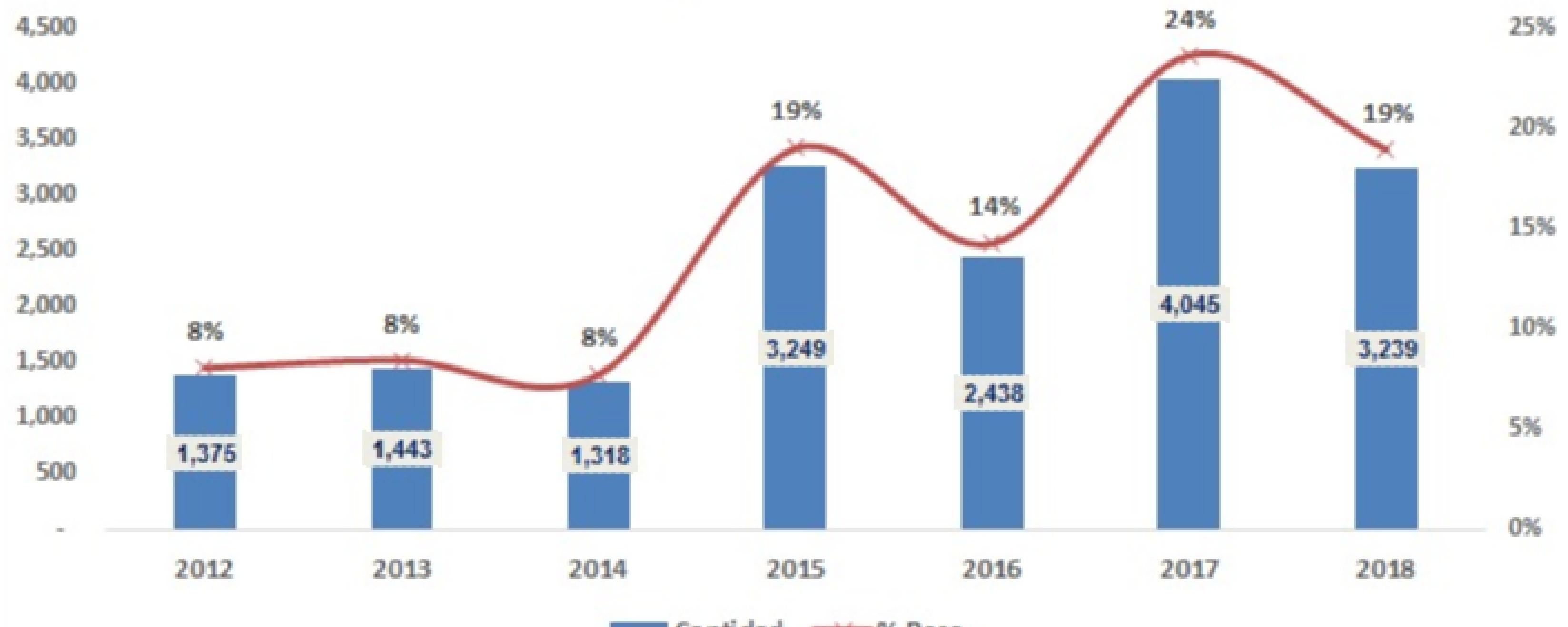
Administracion

D	E	F	G	H	I	J
O	Palabra	Agrupación	- Clase	T	LISTA ENAHO	-
.008	limpieza	Limpieza	Empleo			
.012	administrativa	Administracion	Empleo			
.018	asesor	Asesor	Empleo			
.016	comerci	Comercio	Empleo			
.011	teleoperador	Teleoperador	Empleo			
.014	humano	Recursos Humanos	Empleo			
.010	analista	Analista	Empleo	+		
.009	recurso	Recursos Humanos	Empleo			
.008	profesion	Profesional	Empleo			
.011	asistent	Asistente	Empleo			
.011	practicant	Practicamente	Empleo			
.008	reclutamiento	Recursos Humanos	Empleo			
.008	industri	Industrial	Empleo			
.011	desarrollo	Desarrollador de Software	Empleo			
.010	cajero	Cajero	Empleo			
.015	venta	Vendedor	Empleo			
.011	vendedor	Vendedor	Empleo			
.016	client	Atención al cliente	Empleo			
.016	programador	Desarrollador de Software	Empleo			
.014	jefe	Jefe	Empleo			
.012	venta	Vendedor	Empleo			
.011	servicio	Atención al cliente	Empleo			
.010	analista	Analista	Empleo			

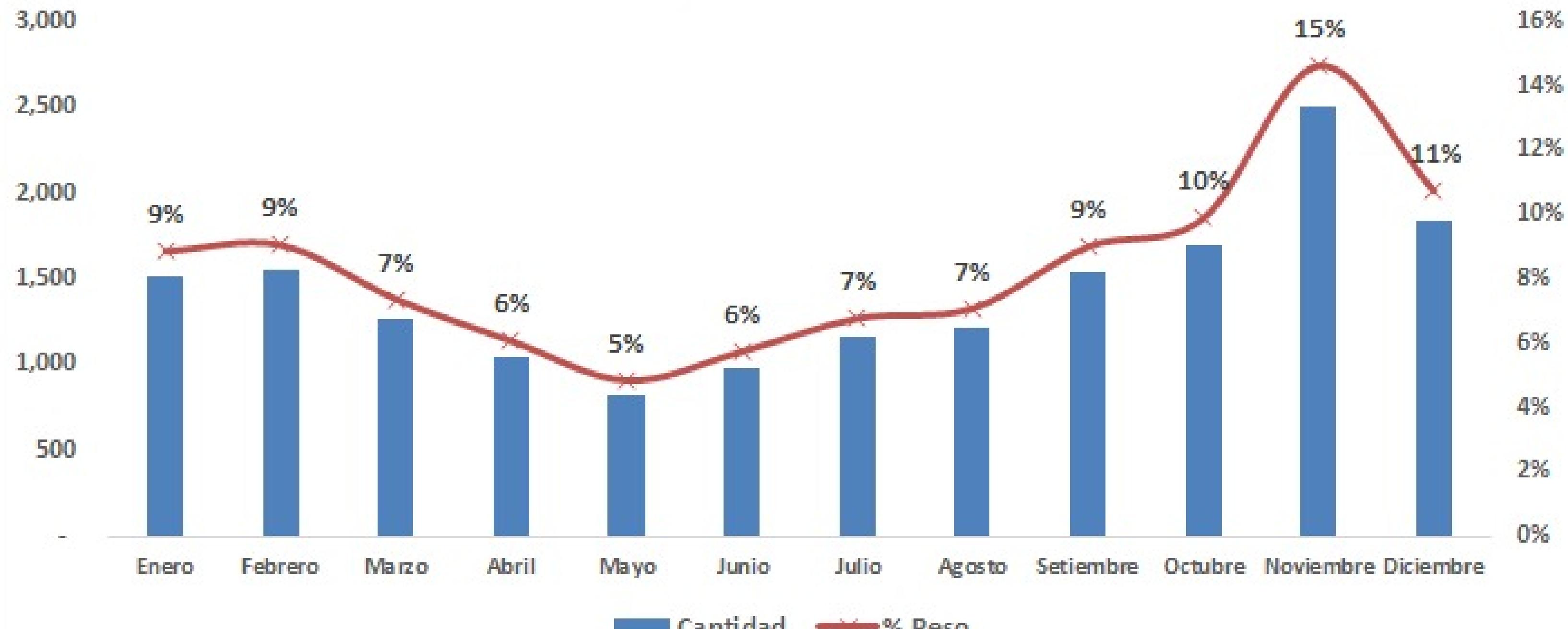
Especialidades más demandadas en Twitter



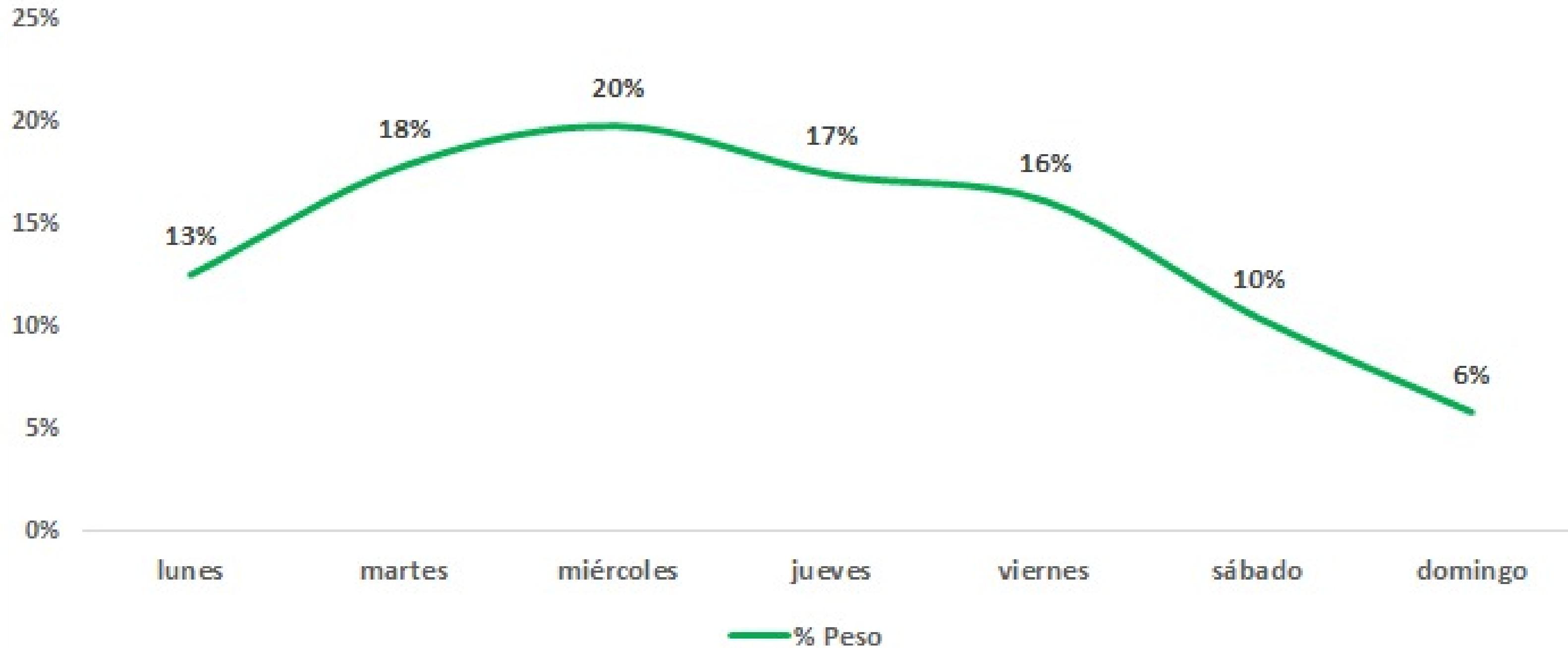
Ofertas laborales publicadas en Twitter x Año



Ofertas laborales publicadas en Twitter x Mes

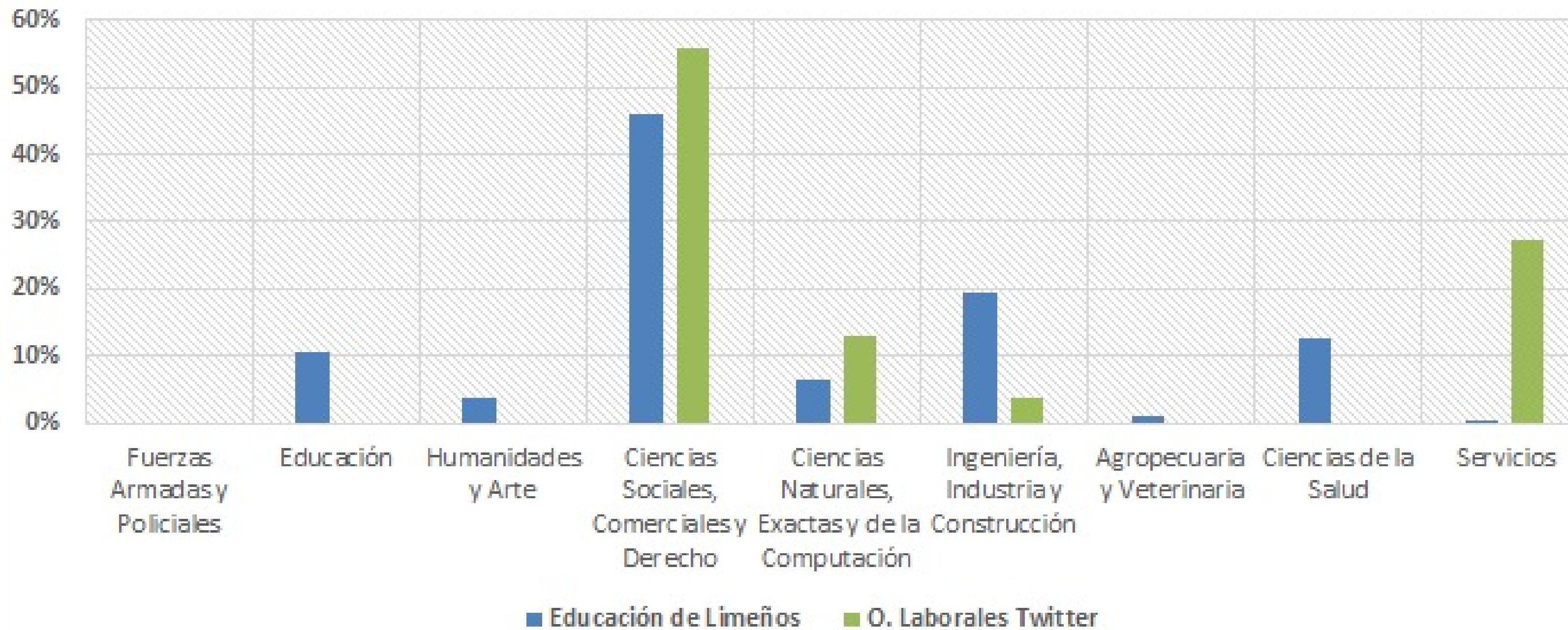


Ofertas laborales publicadas en Twitter x Día

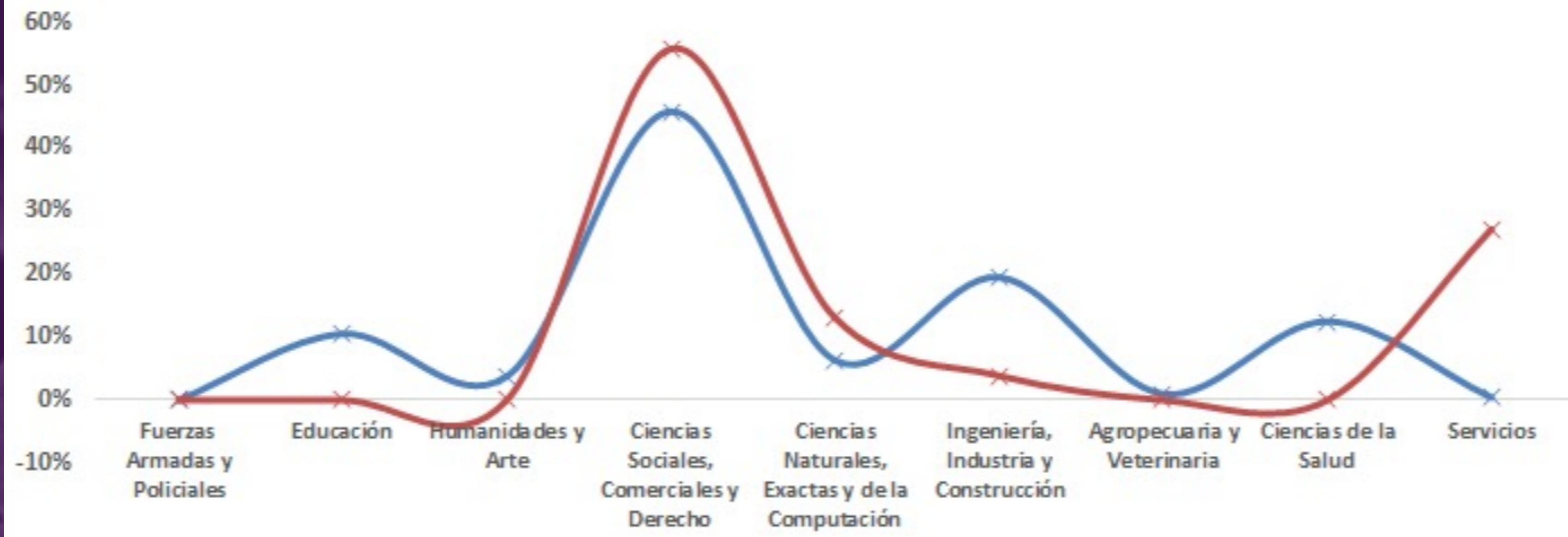


Código	Especialidad	% Peso
3	Ciencias Sociales, Comerciales y Derecho	56%
8	Servicios	27%
4	Ciencias Naturales, Exactas y de la Computación	13%
5	Ingeniería, Industria y Construcción	4%
0	Fuerzas Armadas y Policiales	0%
1	Educación	0%
2	Humanidades y Arte	0%
6	Agropecuaria y Veterinaria	0%
7	Ciencias de la Salud	0%

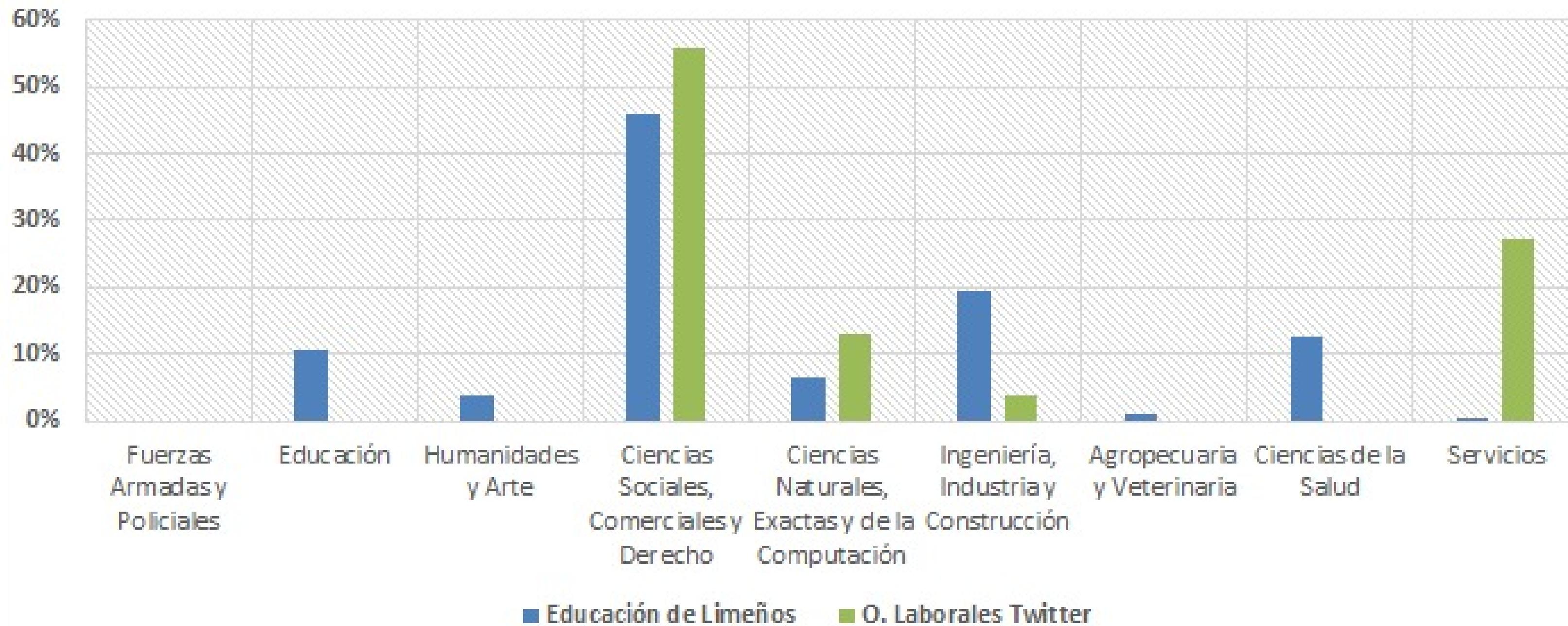
Ofertas laborales publicadas en Twitter VS Educación de Limeños



Ofertas laborales publicadas en Twitter VS Educación de Limeños



Ofertas laborales publicadas en Twitter VS Educación de Limeños



GRACIAS TOTALES





MUCHAS GRACIAS
CTIC

