

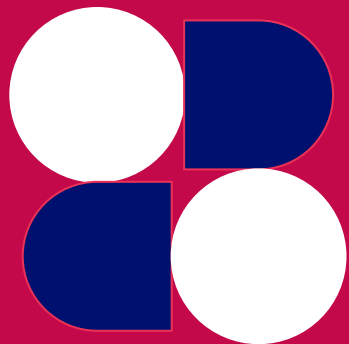


Programa PUCP DATA LITERACY

Formación tecnológica

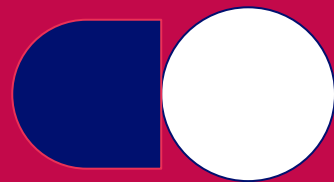
Dirección Académica de Planeamiento y Evaluación
Dirección de Gestión del Talento Humano





SESIÓN 9:

Visualizaciones II





Agenda



Al finalizar la sesión, lograrás:

- Comprender cuándo utilizar un gráfico de tortas, línea, de barras apiladas, un histograma, un gráfico de dispersión y una caja de bigotes.



Contenido:



1

Gráficos para mostrar comparación

2

Gráficos para mostrar composición

3

Gráficos para mostrar distribución

4

Gráficos para mostrar relación



Comparación Gráfico de líneas

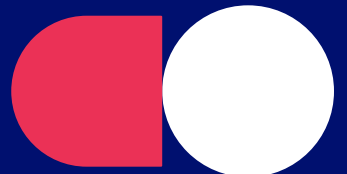




Gráfico de líneas



Número de admitidos en tres universidades seleccionadas 2023-2024



Los gráficos de líneas utilizan una serie de puntos conectados por líneas rectas.



Datos en el tiempo los podemos agregar o desagregar en función de la periodicidad elegida y vamos a poder comparar diferentes elementos

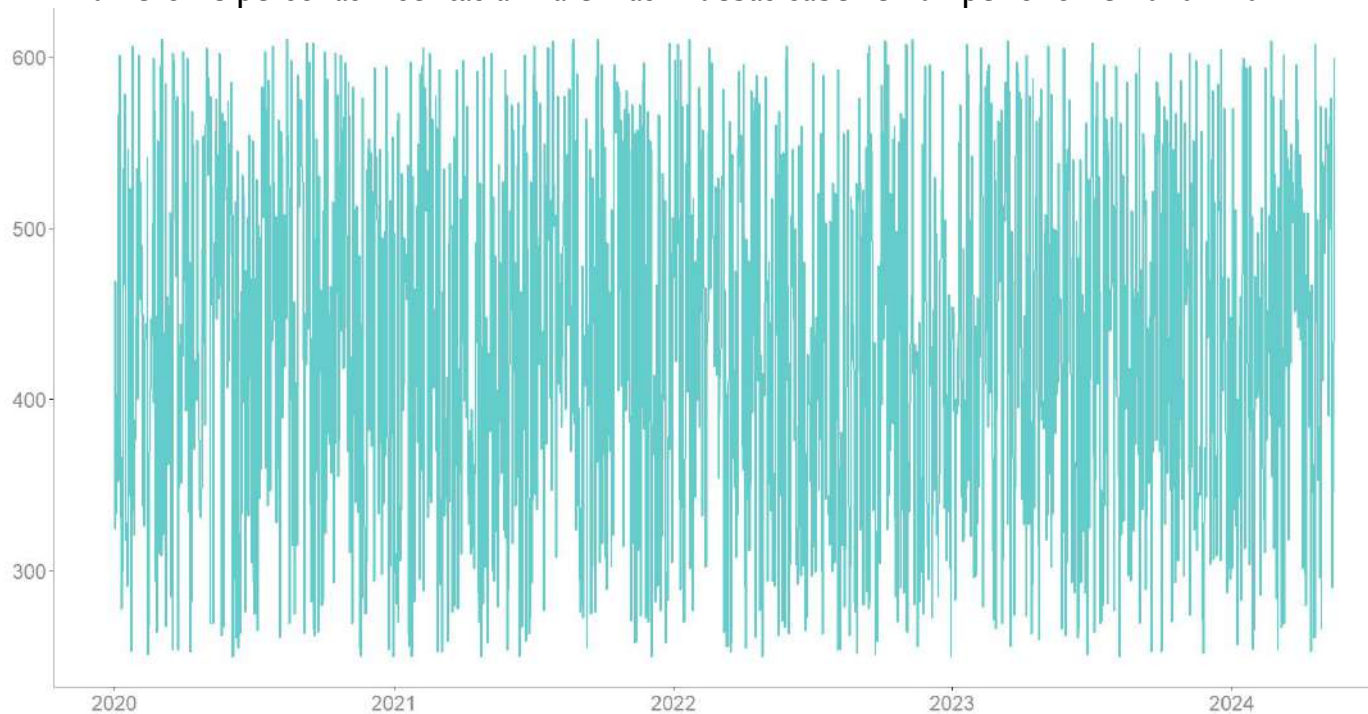
Vamos a ver los “mismos datos” agregados a diferente periodicidad



Recomendación: Datos a diferentes períodos de tiempo- día



Número de personas inscritas al día en las Pruebas Saber en un periodo de 2020 - 2024



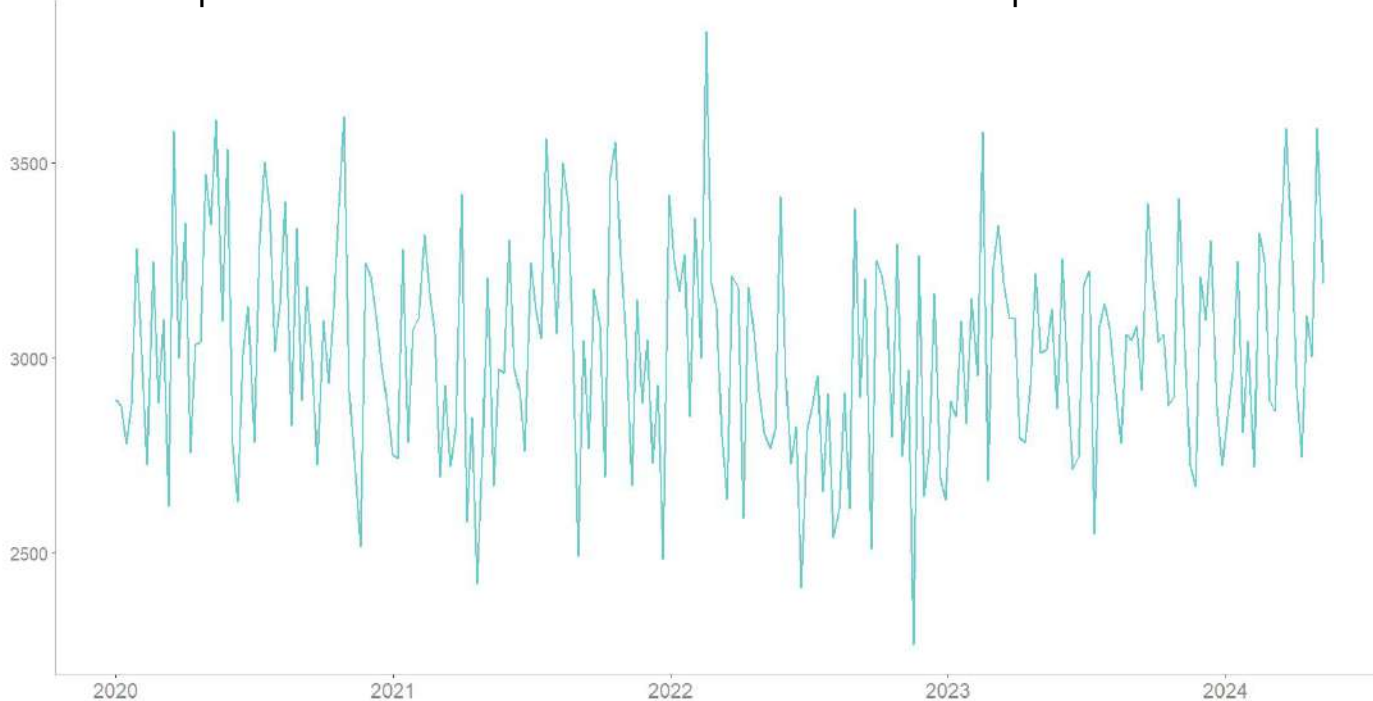
Fuente: Elaboración propia



Recomendación: Datos a diferentes períodos de tiempo- semana



Número de personas inscritas a la semana en las Pruebas Saber en un periodo de 2020 - 2024



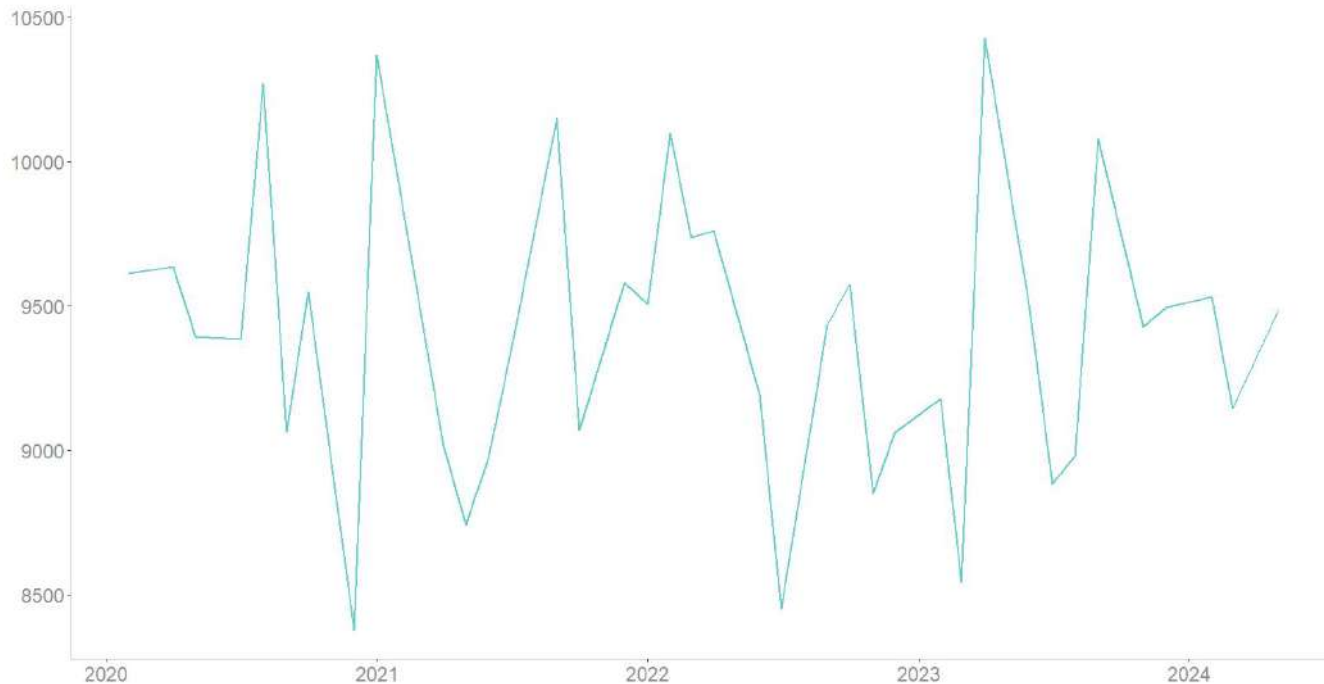
Fuente: Elaboración propia



Recomendación: Datos a diferentes períodos de tiempo- mes



Número de personas inscritas al mes en las Pruebas Saber en un periodo de 2020 - 2024



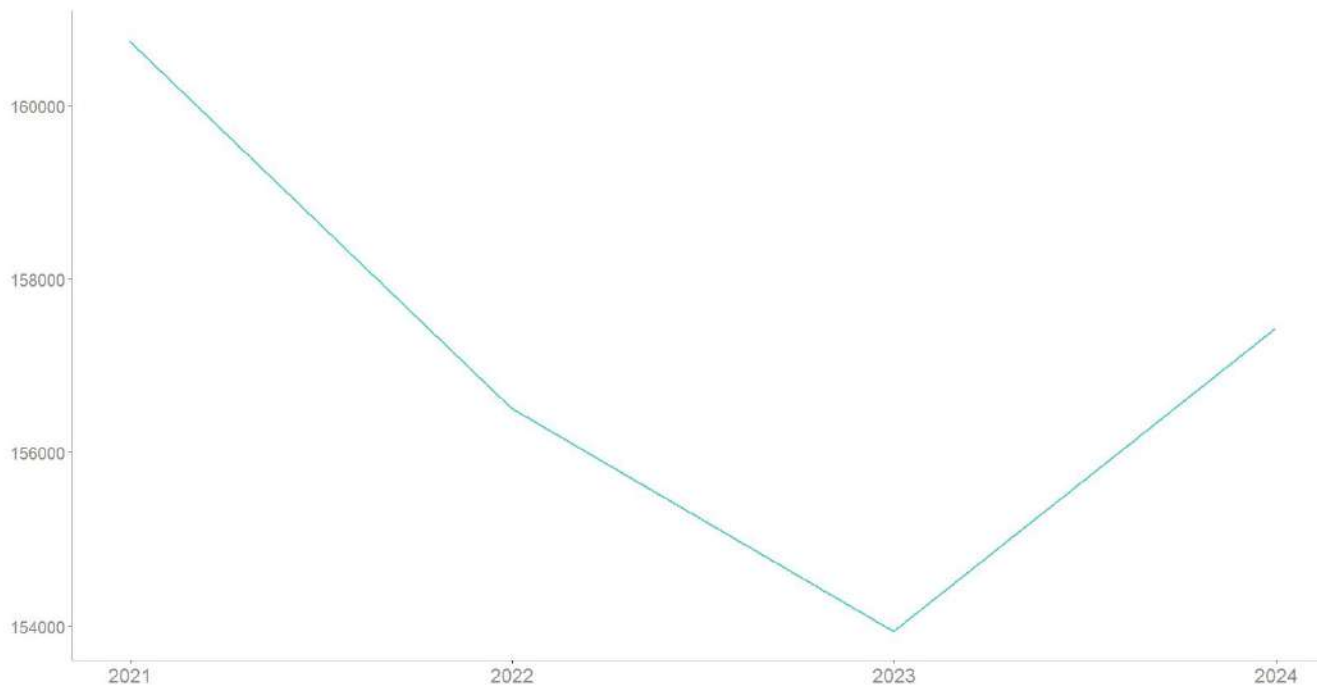
Fuente: Elaboración propia



Recomendación: Datos a diferentes períodos de tiempo- Año



Número de personas inscritas al año en las Pruebas Saber en un periodo de 2020 - 2024



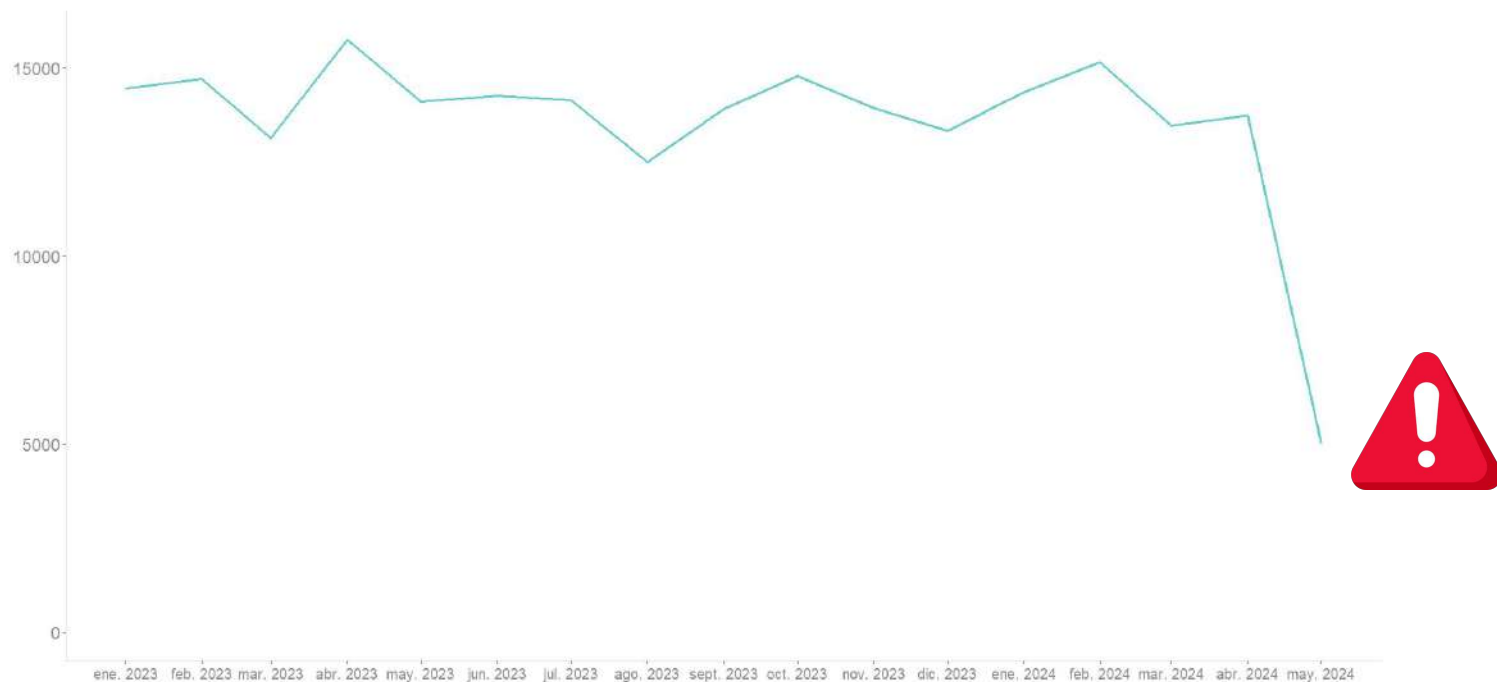
Fuente: Elaboración propia



Recomendación



Número de pagos de derechos para la prueba para Pruebas Saber 2023-2024



Fuente: Elaboración propia



Recomendación



Número de pagos de derechos para la prueba para Pruebas Saber 2023-2024



Fuente: Elaboración propia

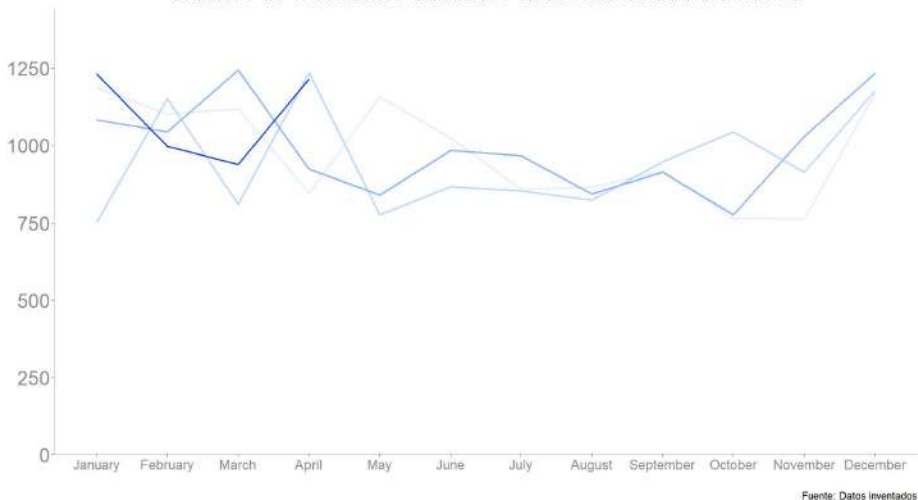


Gráfico de líneas



Las líneas conectadas son más expresivas del cambio a lo largo del tiempo que las barras, dado que representan visualmente la continuidad vs. las comparaciones discretas.

Número de estudiantes admitidos en la Universidad del Cauca



Número de estudiantes admitidos en la Universidad del Cauca

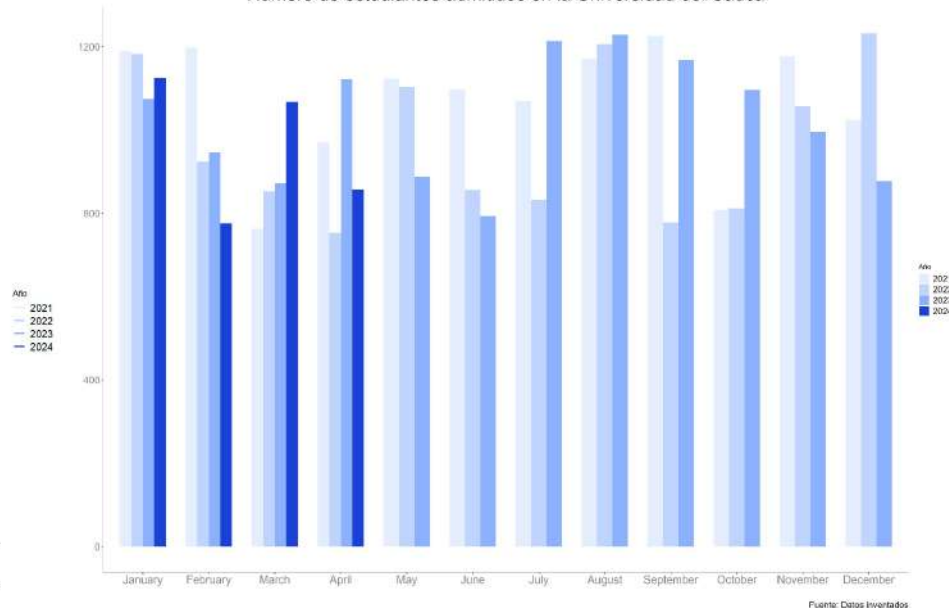
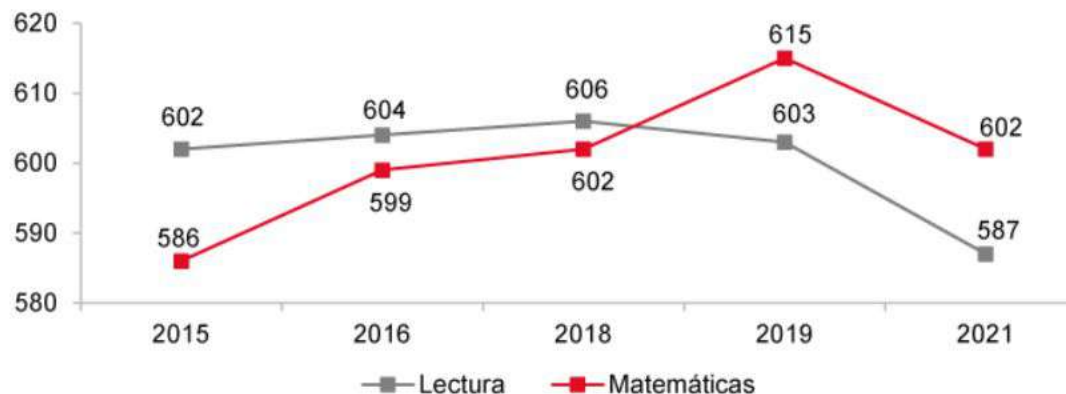




Gráfico de líneas

Evolución de resultados en evaluaciones de alumnos de 2.º grado de secundaria (2015-2021)



Fuente: Minedu. Elaboración: ComexPerú.

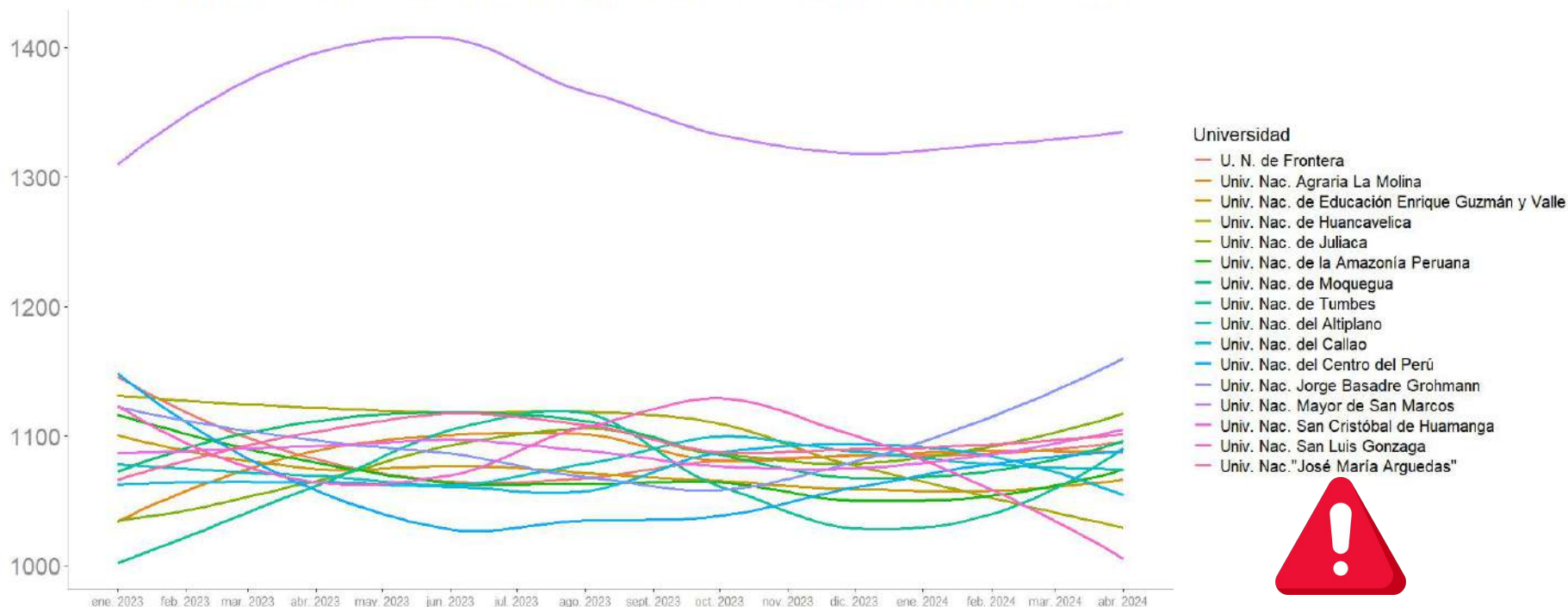




Gráfico de líneas



Número de estudiantes admitidos en 16 universidades seleccionadas en Perú 2023-2024



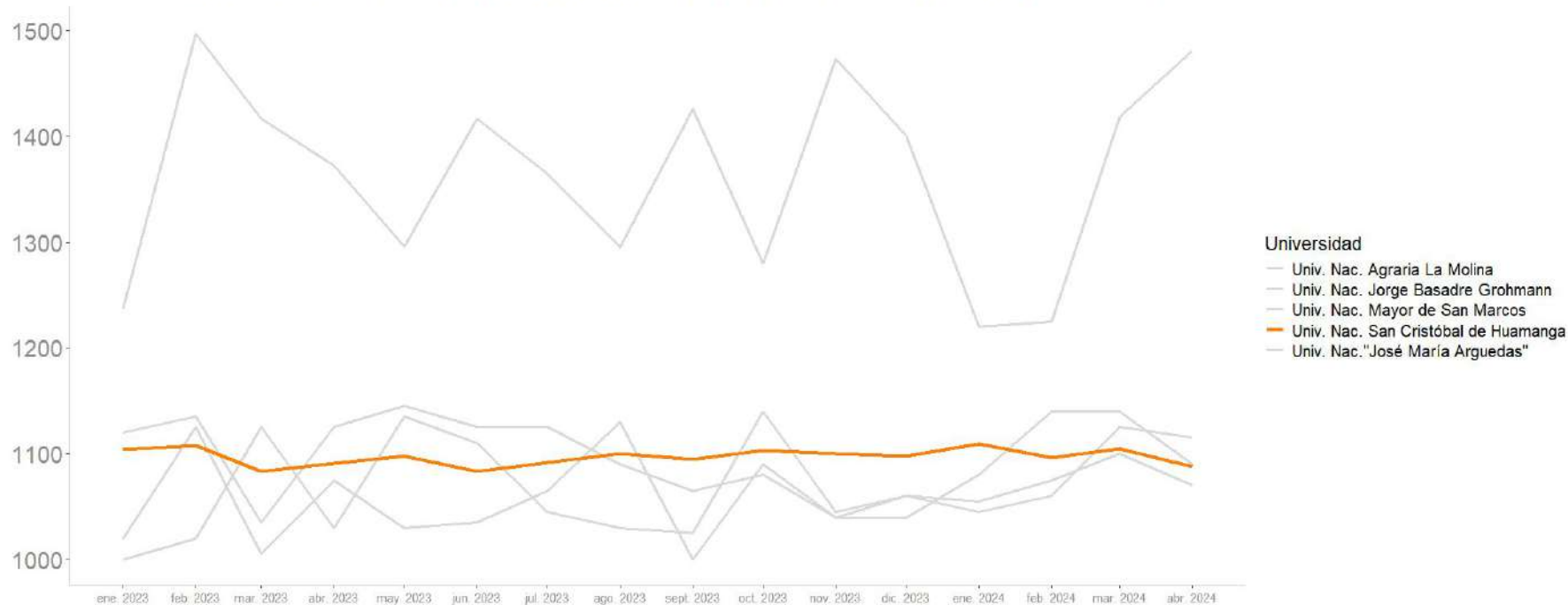
Fuente: Elaboración propia



Gráfico de líneas



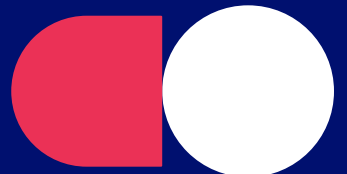
Número de estudiantes admitidos en 5 universidades seleccionadas en Perú 2023-2024



Fuente: Elaboración propia



Composición





A menudo, nuestros datos pueden agruparse en niveles de categoría que, al combinarse, forman una unidad o un total significativo.





Composición



1.

Gráfico de Barras apiladas

2.

Gráfico circular o gráfico de tortas



Gráfico de Barras Apiladas



Gráfico de Barras Apiladas



Porcentaje de estudiantes admitidos de tres Universidades en Colombia

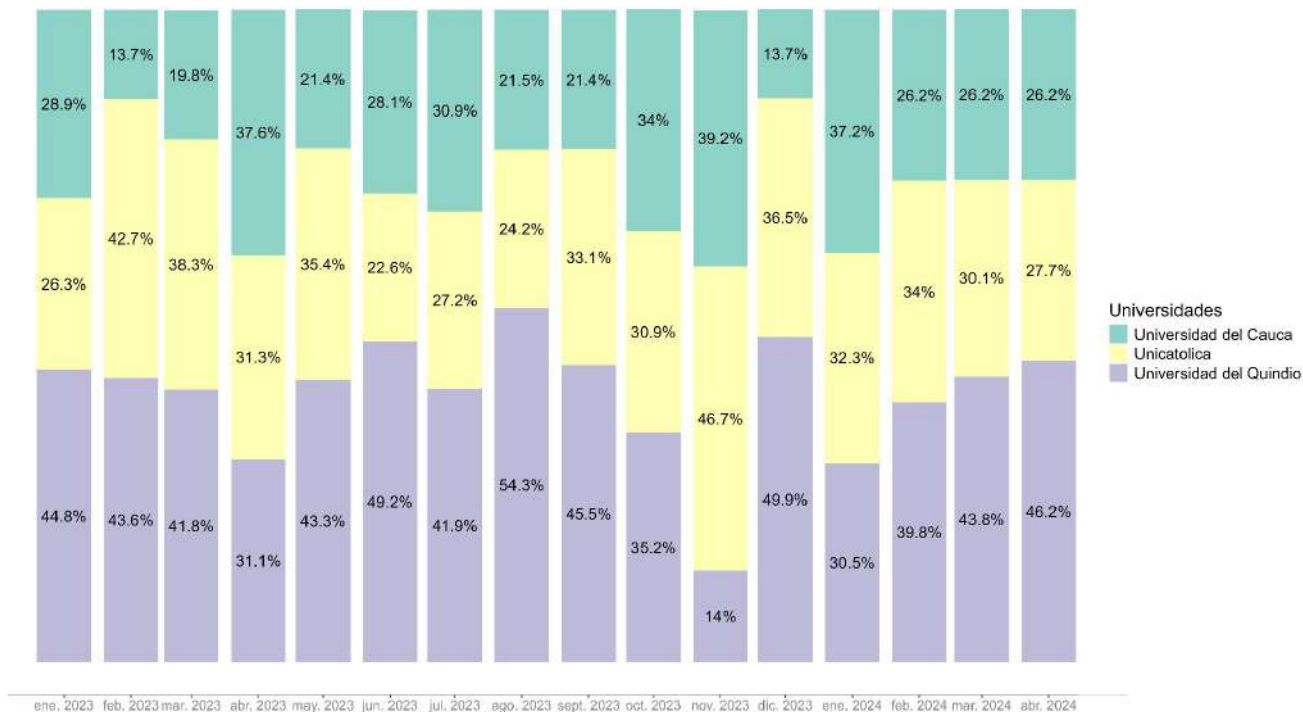




Gráfico de Barras Apiladas



Número de estudiantes admitidos de tres Universidades en Colombia



Fuente: Datos inventados

Porcentaje de estudiantes admitidos de tres Universidades en Colombia



Fuente: Datos inventados

¿Qué es lo más importante al crear un gráfico de barras apiladas?



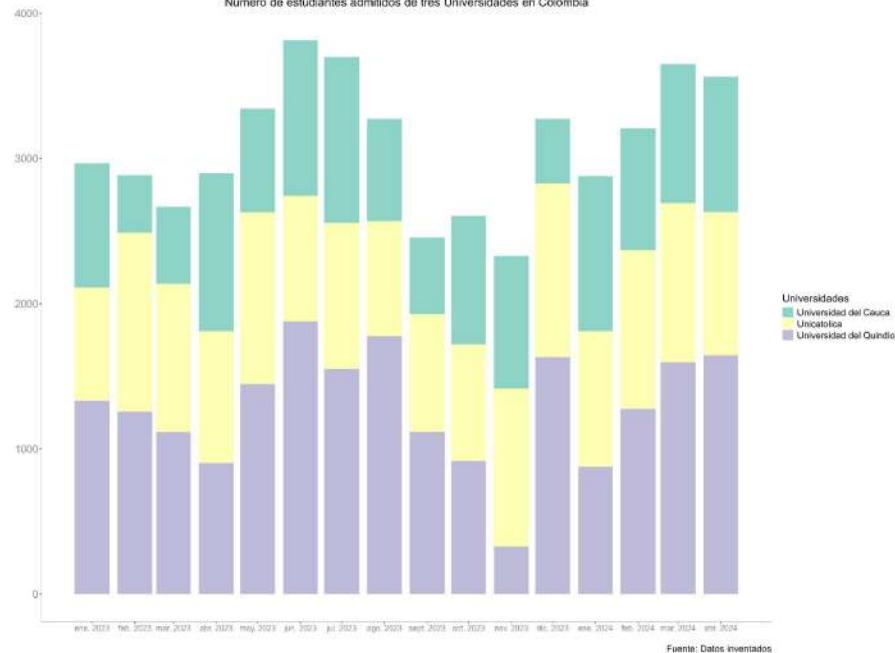
Gráfico de Barras Apiladas



Número de estudiantes admitidos de tres Universidades en Colombia



Número de estudiantes admitidos de tres Universidades en Colombia



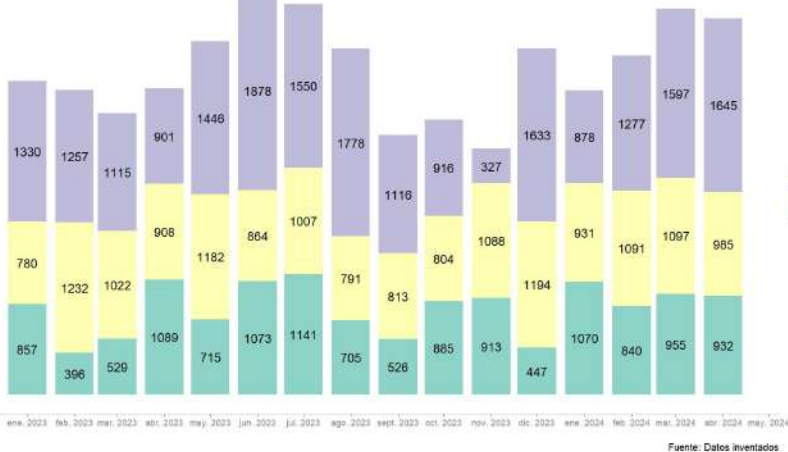
Los segmentos apilados no comparten una línea base común, por lo que puede ser más desafiante comparar con precisión los tamaños relativos de cada segmento individual.



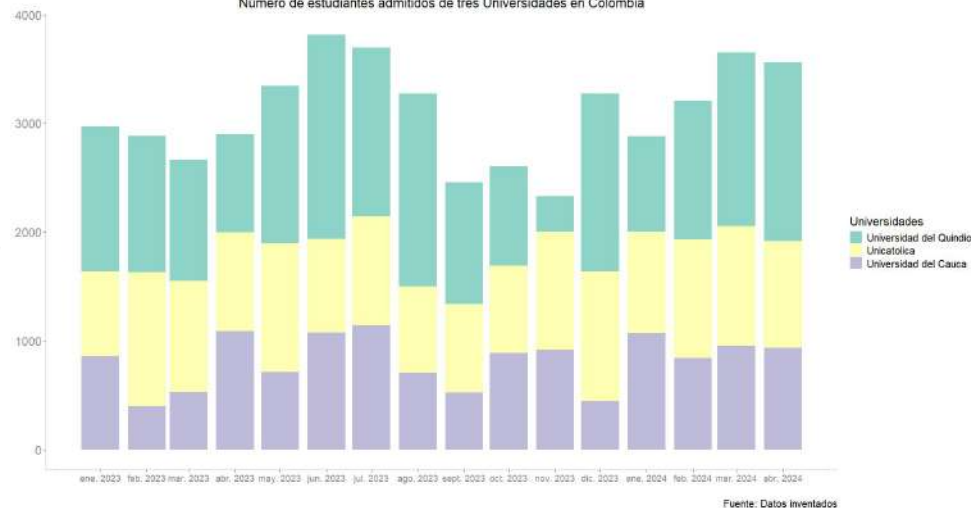
Gráfico de Barras Apiladas



Número de estudiantes admitidos de tres Universidades en Colombia



Número de estudiantes admitidos de tres Universidades en Colombia



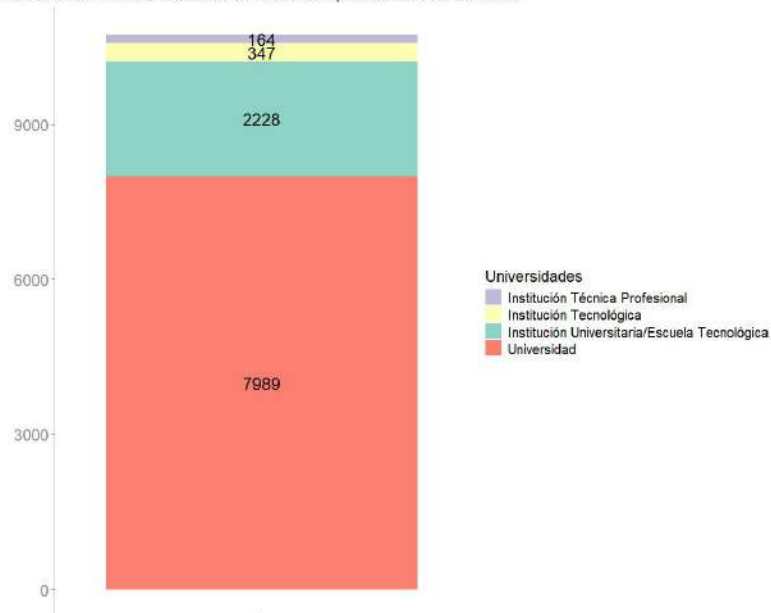
Los segmentos apilados no comparten una línea base común, por lo que puede ser más desafiante comparar con precisión los tamaños relativos de cada segmento individual.



Si es necesario hacer comparaciones altamente precisas y exactas entre diferentes partes del todo, un gráfico de barras estándar puede ser la mejor opción.

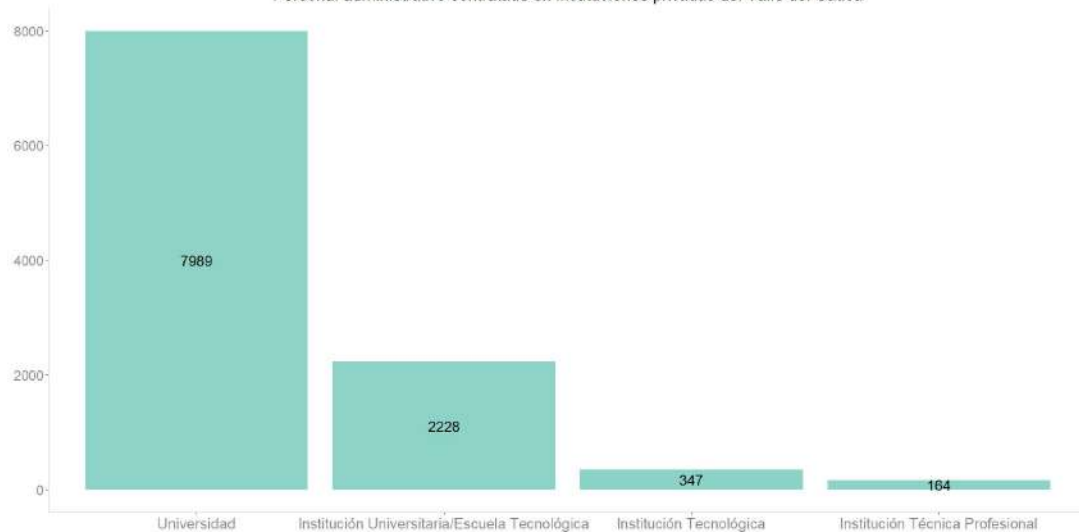


Personal administrativo contratado en instituciones privadas del Valle del Cauca



Fuente: ministerio de educación nacional de colombia

Personal administrativo contratado en instituciones privadas del Valle del Cauca



Fuente: ministerio de educación nacional de colombia



**Gráfico circular o de
tortas**



Un gráfico circular es simplemente un gráfico de barras apiladas creado utilizando coordenadas polares en lugar de rectangulares.

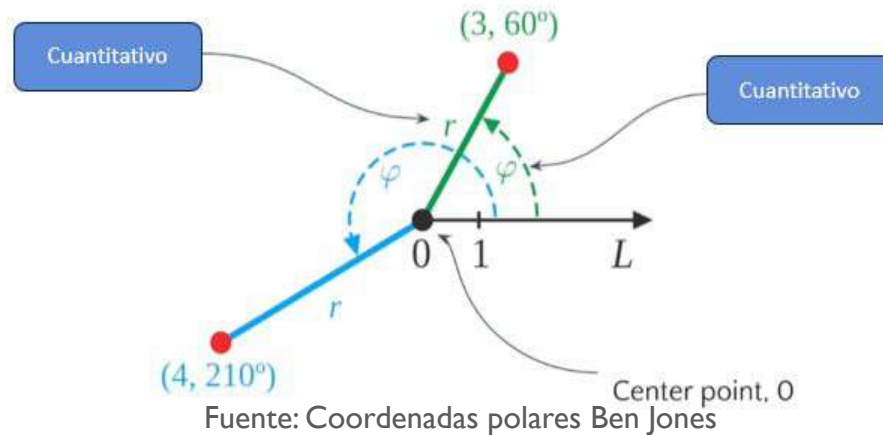
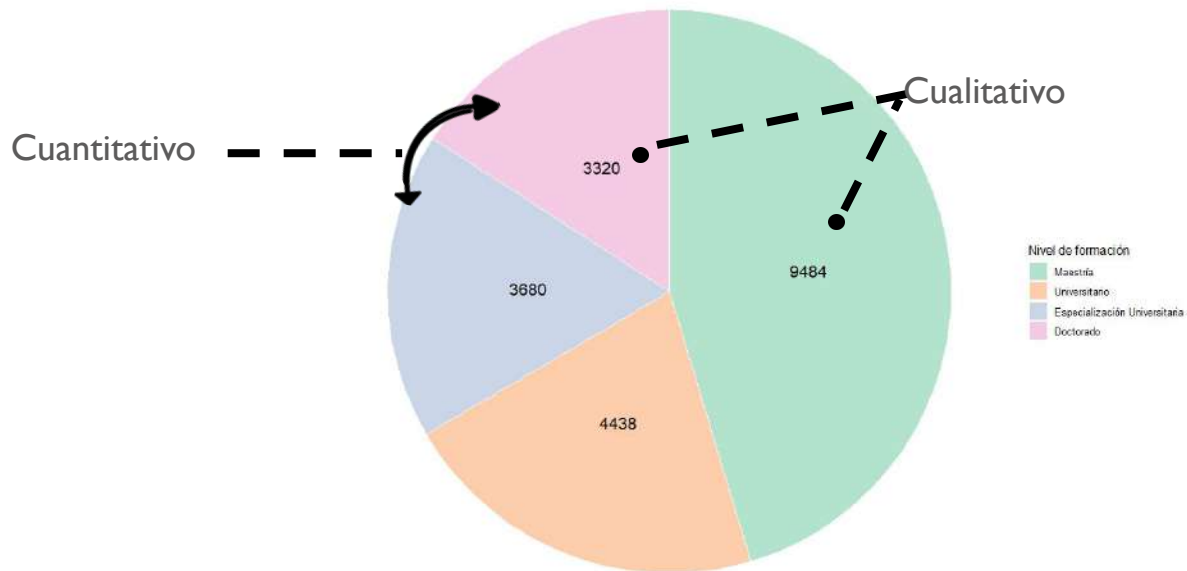




Gráfico circular

Nivel de formación de docentes contratados en Universidades del Valle del Cauca



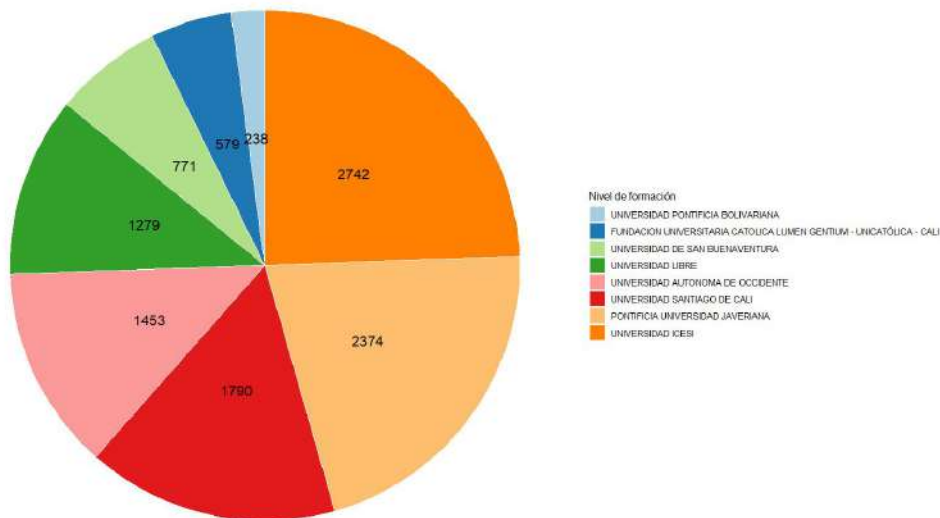
Fuente: Ministerio de educación nacional de colombia



Muchos componentes pueden ser un problema particular, debido a que puede resultar visualmente abrumador para algunas personas.



Número de docentes contratados en Universidades privadas seleccionadas en el Valle del Cauca 2022



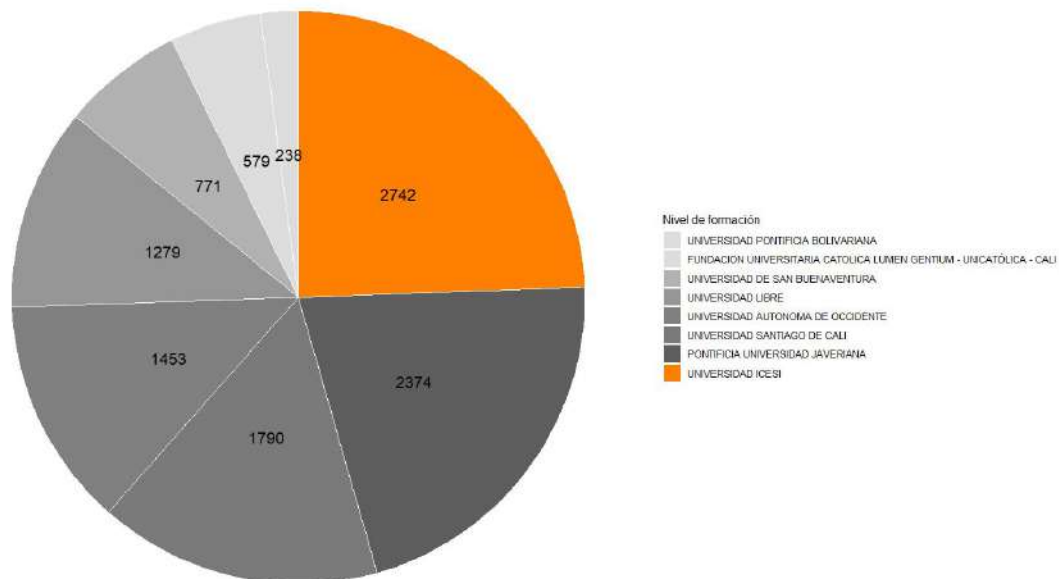
Fuente: Ministerio de educación nacional de Colombia



Pueden haber excepciones



Número de docentes contratados en Universidades seleccionadas privadas en el Valle del Cauca

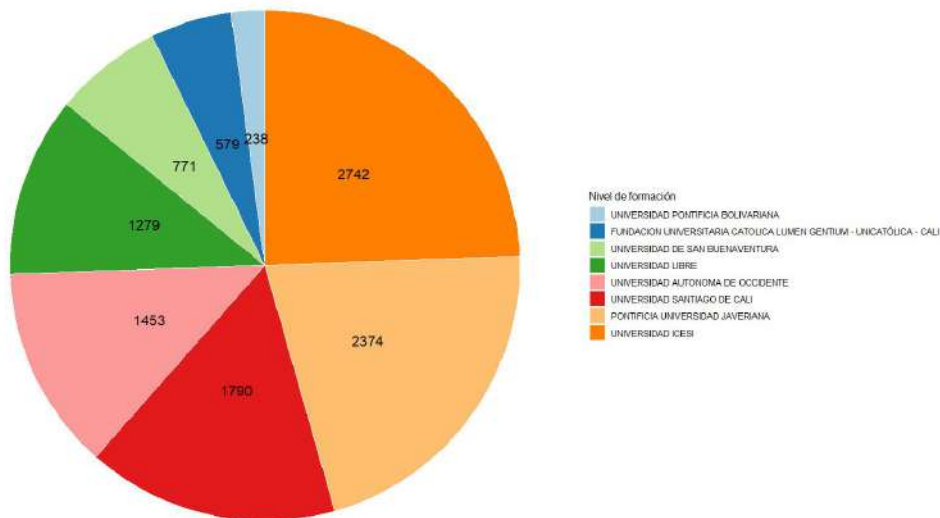




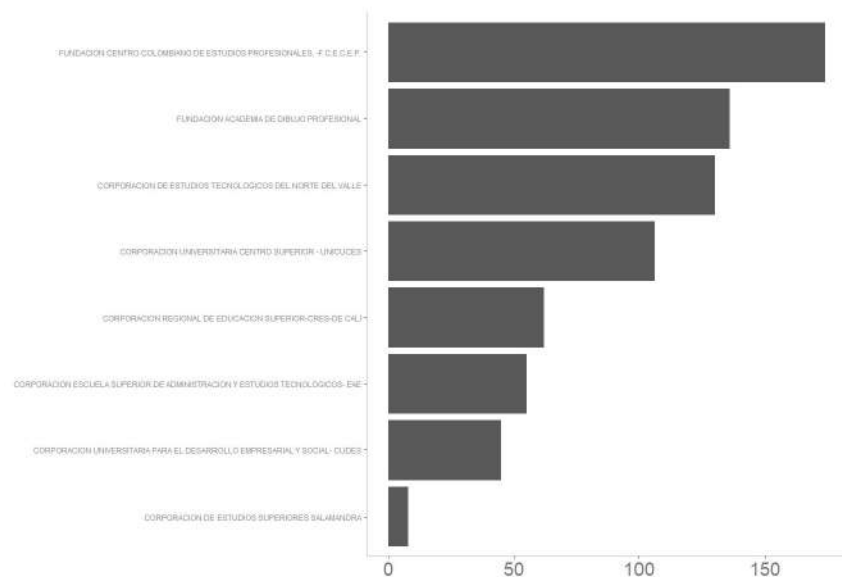
Muchos componentes pueden ser un problema particular, debido a que puede resultar visualmente abrumador para algunas personas.



Número de docentes contratados en Universidades privadas seleccionadas en el Valle del Cauca 2022



Fuente: Ministerio de educación nacional de colombia



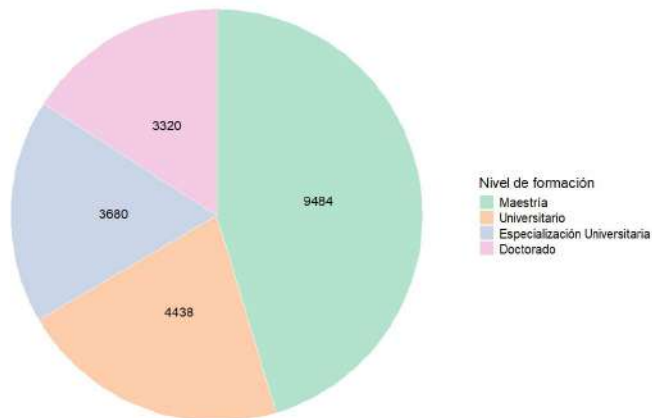
Fuente: Ministerio de educación nacional de colombia



Se debe iniciar en el ángulo de 90 grados, ordenando las secciones de mayor a menor proporción.

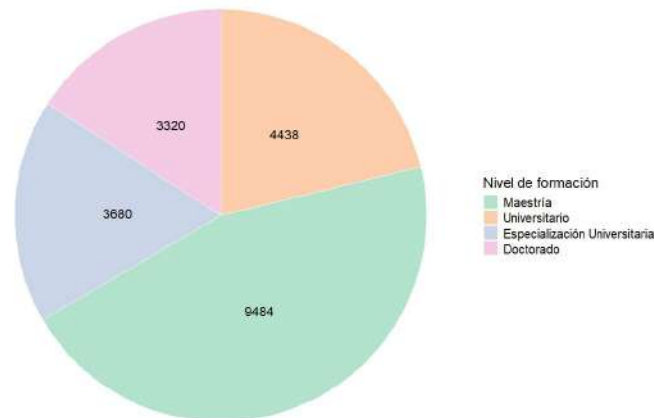


Nivel de formación de docentes contratados en Universidades del Valle del Cauca



Fuente: Ministerio de educación nacional de colombia

Nivel de formación de docentes contratados en Universidades del Valle del Cauca



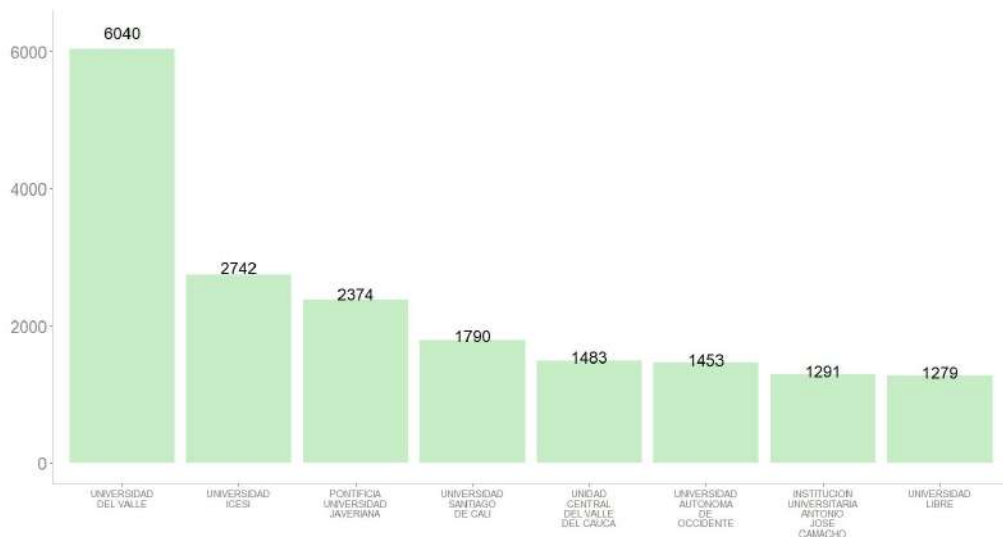
Fuente: Ministerio de educación nacional de colombia



Los gráficos circulares pueden dificultar la comparación precisa de los segmentos de área y si hacer tales comparaciones es importante, podría ser recomendable usar un gráfico de barras.



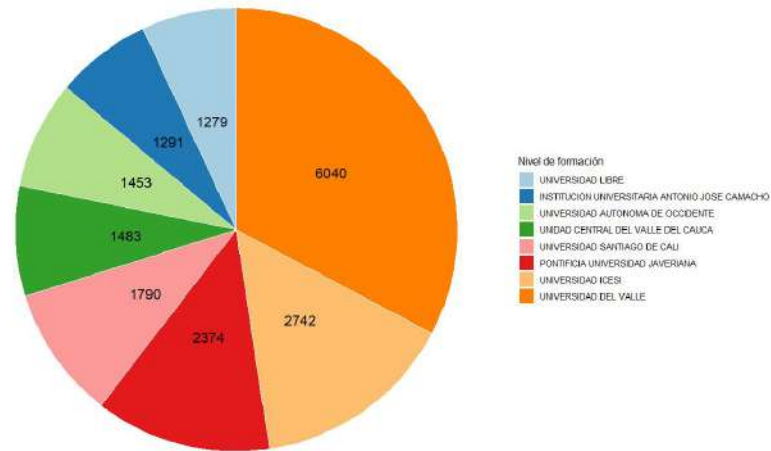
Número de docentes contratados en Universidades privadas seleccionadas en el Valle del Cauca 2022



Fuente: Ministerio de educación nacional de colombia



Número de docentes contratados en Universidades privadas seleccionadas en el Valle del Cauca 2022



Fuente: Ministerio de educación nacional de colombia



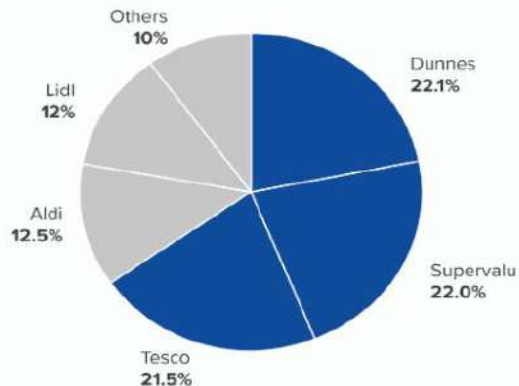
Podríamos hacer una excepción



Similarity story

Dunnes, Supervalu and Tesco all have an equal share of the Irish supermarket sector

Grocery market share Q4 2020

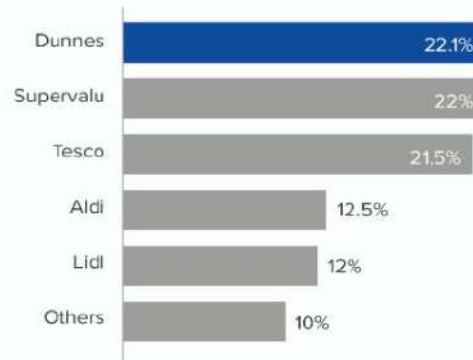


Source: Kantar Worldpanel
<https://www.kantarworldpanel.com/en/grocery-market-share/great-britain/snapshot>

Difference story

Dunnes edges ahead of Supervalu and is Ireland's favourite supermarket again

Grocery market share Q4 2020



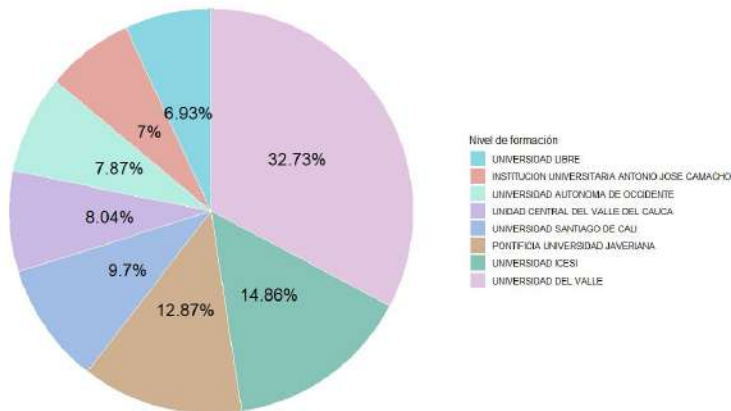
Source: Kantar Worldpanel
<https://www.kantarworldpanel.com/en/grocery-market-share/great-britain/snapshot>



Dado que el gráfico forma un círculo completo, los sectores deben ser mutuamente excluyentes y colectivamente exhaustivos. Es decir, el total de los segmentos no debe sumar más del 100% y los segmentos no deben tener partes faltantes del todo.



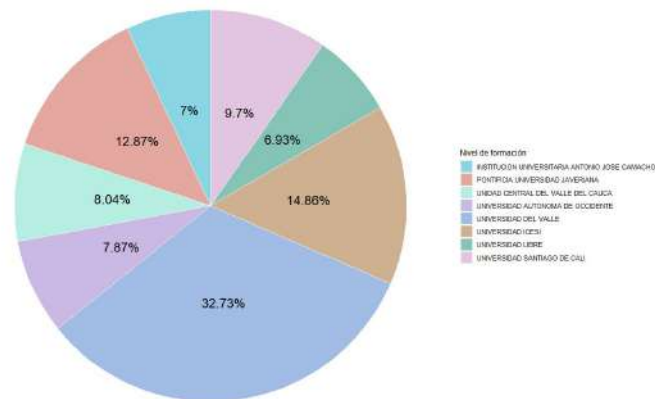
Porcentaje de docentes contratados en Universidades seleccionadas en el Valle del Cauca 2022



Fuente: Ministerio de educación nacional de colombia



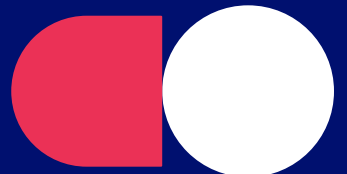
Porcentaje de docentes contratados en Universidades seleccionadas en el Valle del Cauca 2022



Fuente: Ministerio de educación nacional de colombia



Actividad grupal





Actividad



Grupos: Todos pueden participar

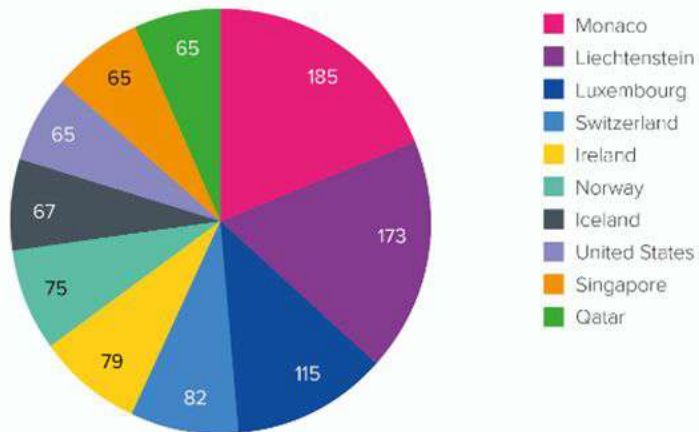
Descripción: Se mostrarán gráficos y tienen que identificar cómo los podríamos mejorar



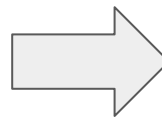
Actividad

The wealth of nations

Top 10 richest countries
(GDP per person in US\$ 000s, 2019)

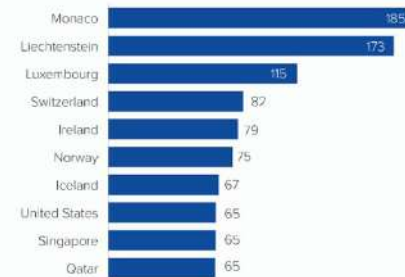


Source: World Bank



The wealth of nations

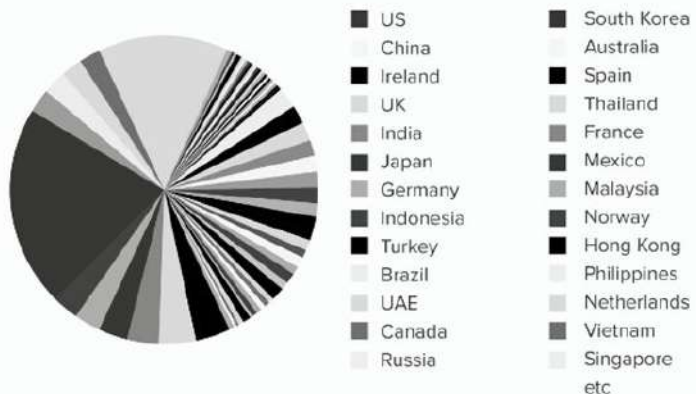
Top 10 richest countries
(GDP per person in US\$ 000s, 2019)



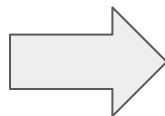
Source: World Bank



Airline passengers by nationality, 2017 (%)

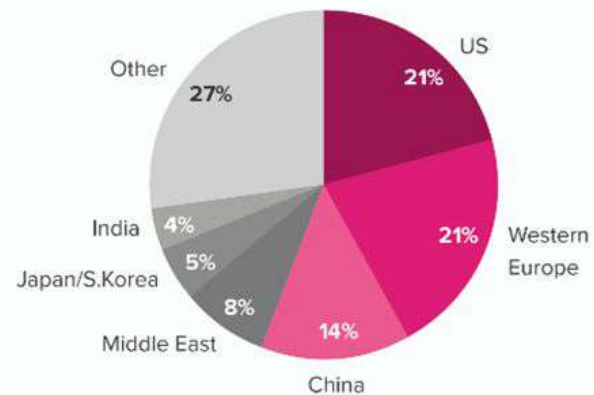


Source: World Bank



Most airline passengers are from US, Western Europe or China

Passengers by nationality, 2017 (%)



Source: World Bank

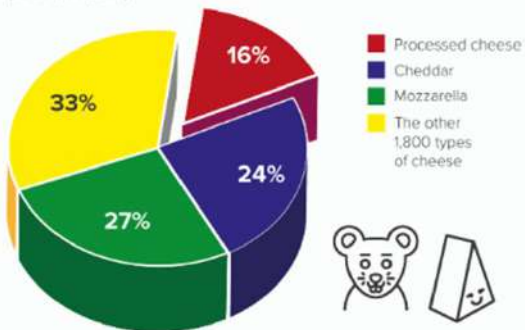


Actividad

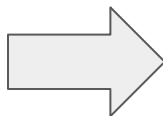
Three key cheeses

US consumers turn their noses up at most varieties of cheese

Percentage of total cheese consumed by type
(lbs per person, US, 2017)



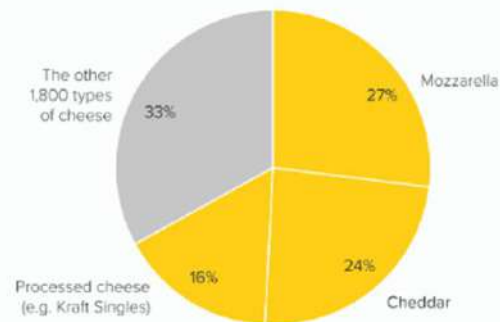
Source: USDA ERS. Mouse icon by Jirapark, cheese by LAFS via the Noun Project



Three key cheeses

US consumers turn their noses up at most varieties of cheese

Percentage of total cheese consumed by type
(lbs per person, US, 2017)



Source: USDA ERS

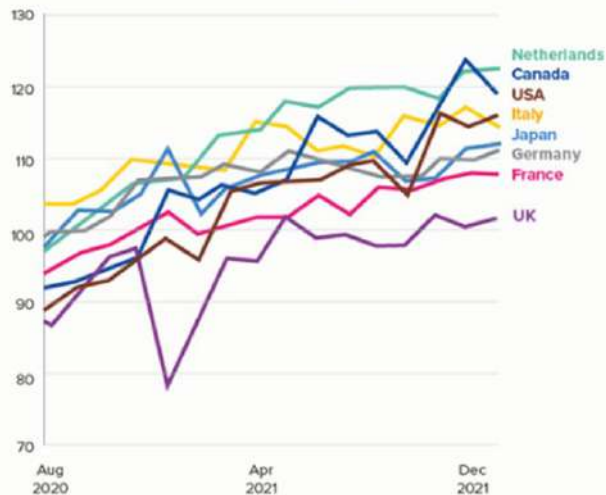


Actividad

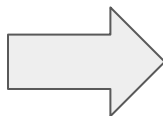
Rough trade

Post-pandemic, the UK has exported fewer goods than similar economies, probably because of Brexit

Goods exports after pandemic (Dec 2019 = 100)



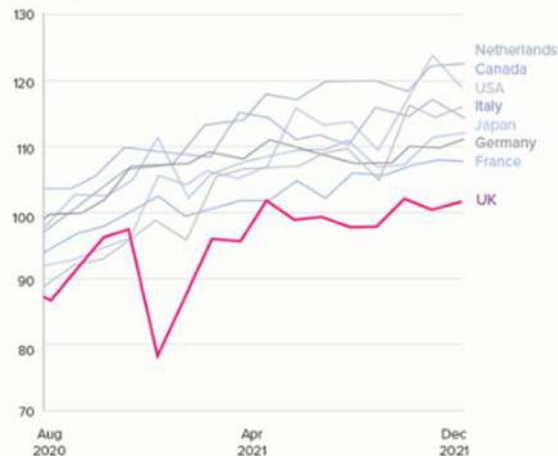
Source: Deloitte



Rough trade

Post-pandemic, the UK has exported fewer goods than similar economies, probably because of Brexit

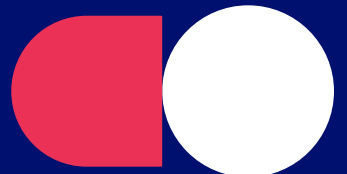
Goods exports after pandemic (Dec 2019 = 100)



Source: Deloitte



Distribución





Composición



1.

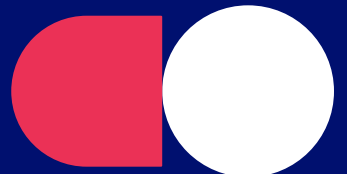
Gráfico de histograma

2.

Gráfico de caja de bigotes

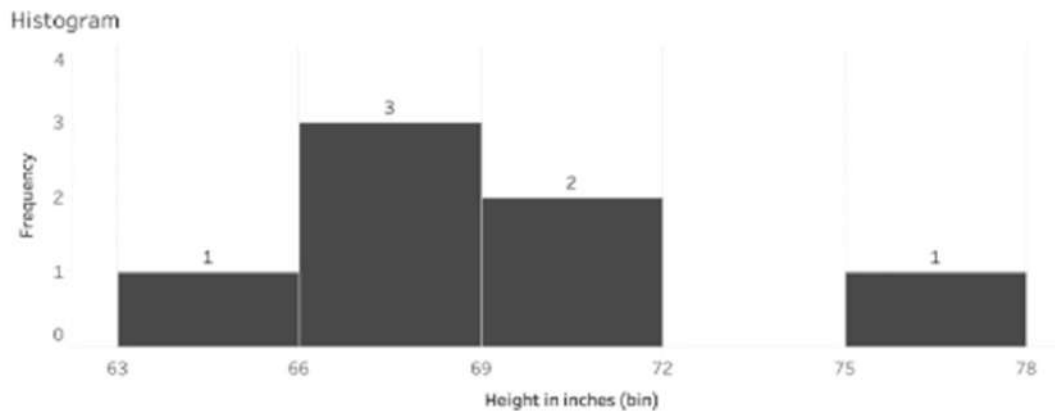


Histograma





Cuantitativo



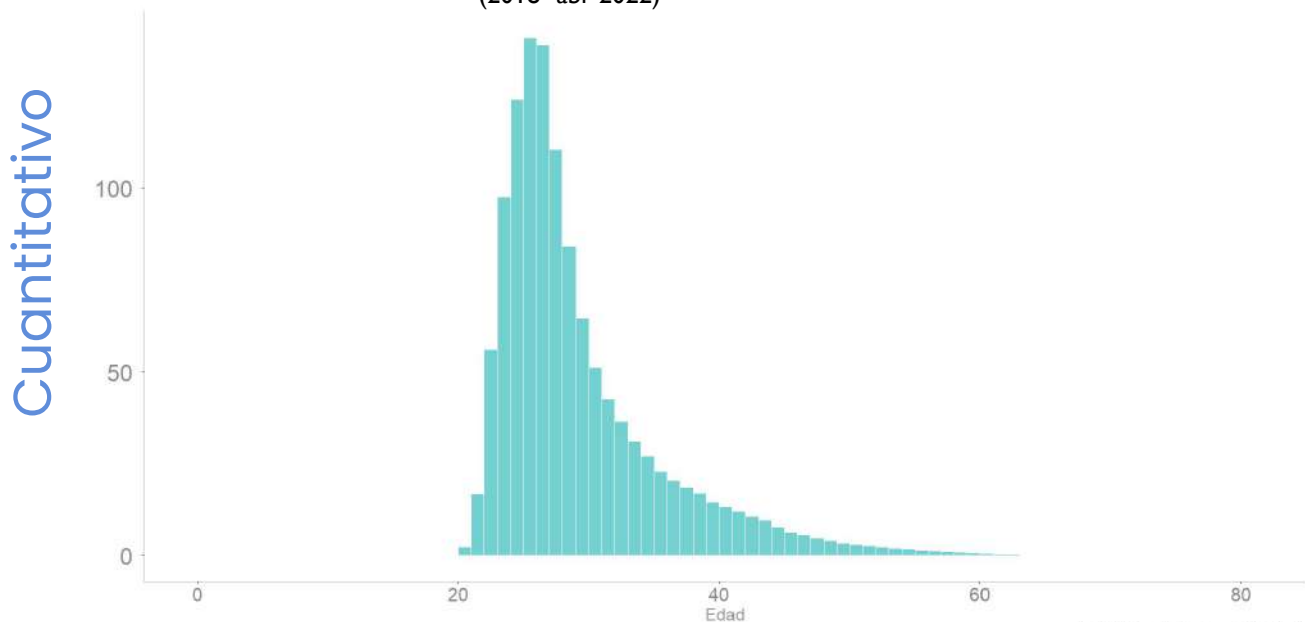
Cuantitativo

Un histograma proporciona una representación visual de la distribución de los datos.



Histograma

Distribución de la edad de las personas que presentaron las pruebas saber pro
(2018- abr 2022)



Fuente: Ministerio de educación nacional de Colombia

Cuantitativo

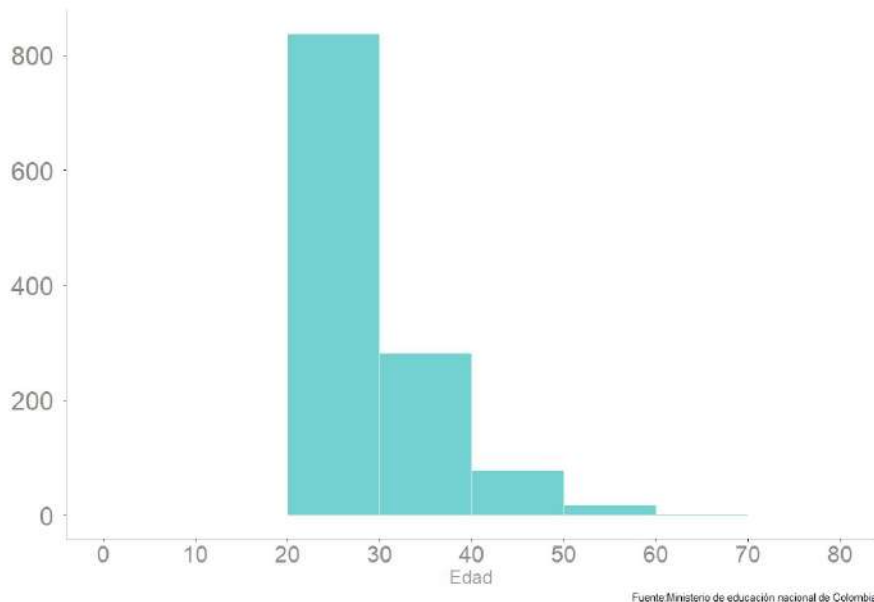
Un histograma proporciona una representación visual de la distribución de los datos.



Histograma

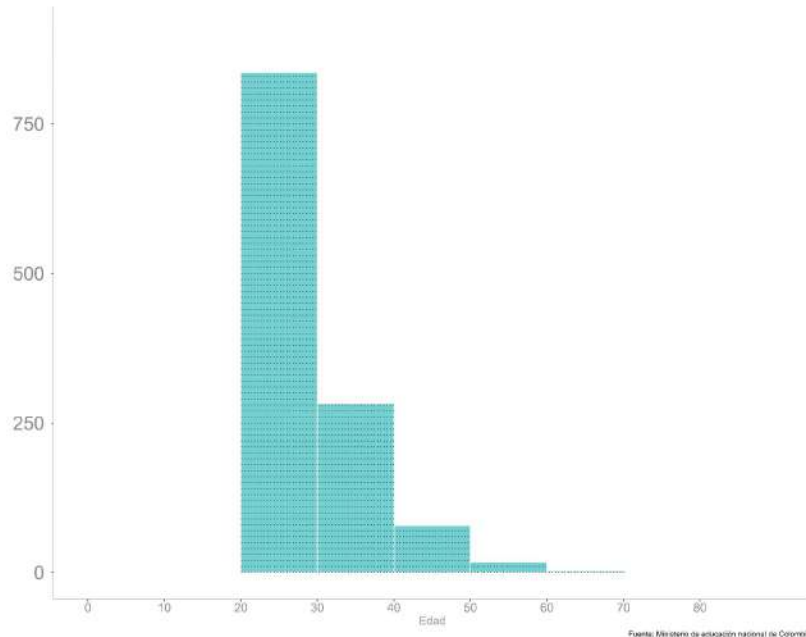


Distribución (En Miles) de la edad de las personas que presentaron las pruebas saber pro (2018- abr 2022)



Distribución (En Miles) de la edad de las personas que presentaron las pruebas saber pro (2018- abr 2022)

Numero de contenedores por cada 10000



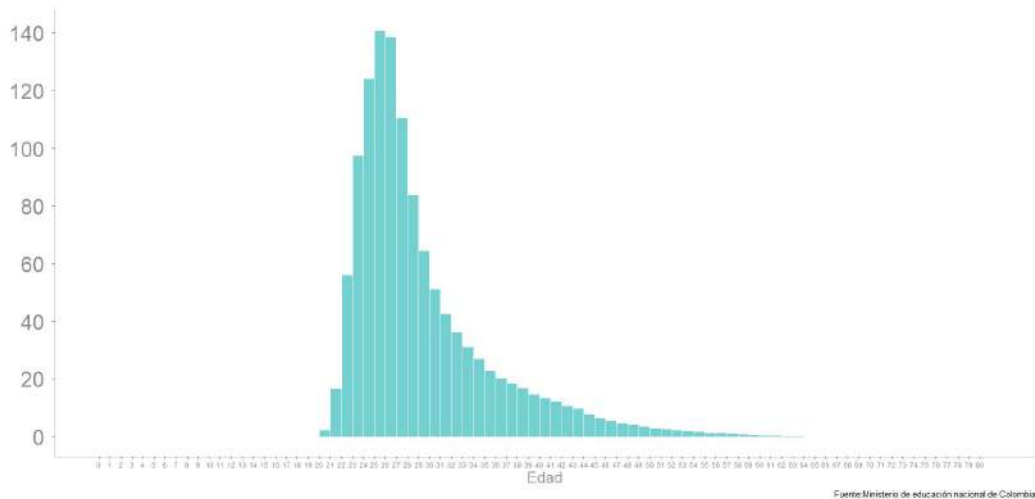
Un histograma proporciona una representación visual de la distribución de los datos.



Histograma

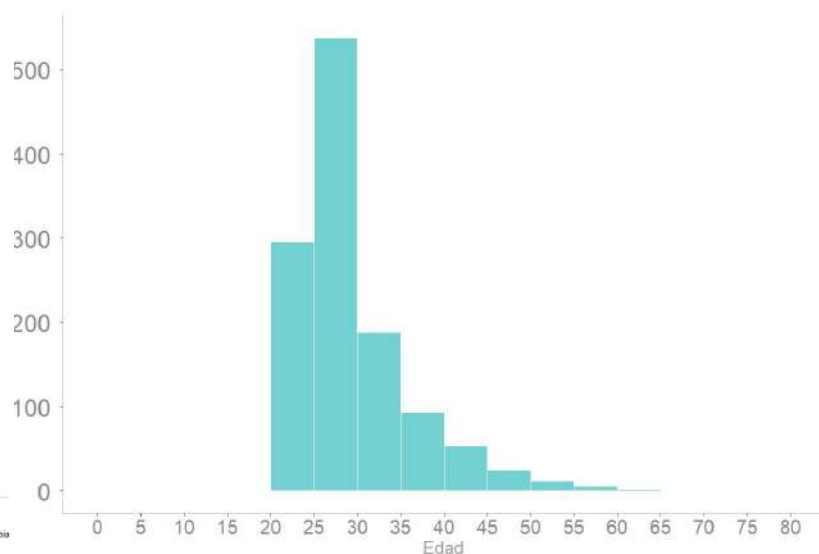


Distribución de la edad de las personas que presentaron las pruebas saber pro (2018- abr 2022)



Fuente: Ministerio de educación nacional de Colombia

Distribución de la edad de las personas que presentaron las pruebas saber pro (2018- abr 2022)



Fuente: Ministerio de educación nacional de Colombia

Los tamaños de los contenedores de un histograma pueden modificarse.

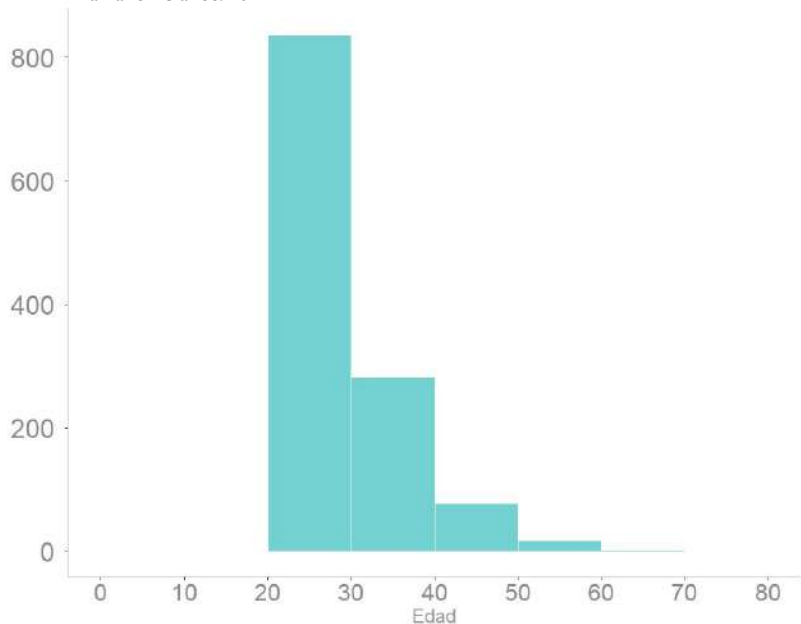


Histograma



Distribución de la edad de las personas que presentaron las pruebas saber pro (2018- abr 2022)

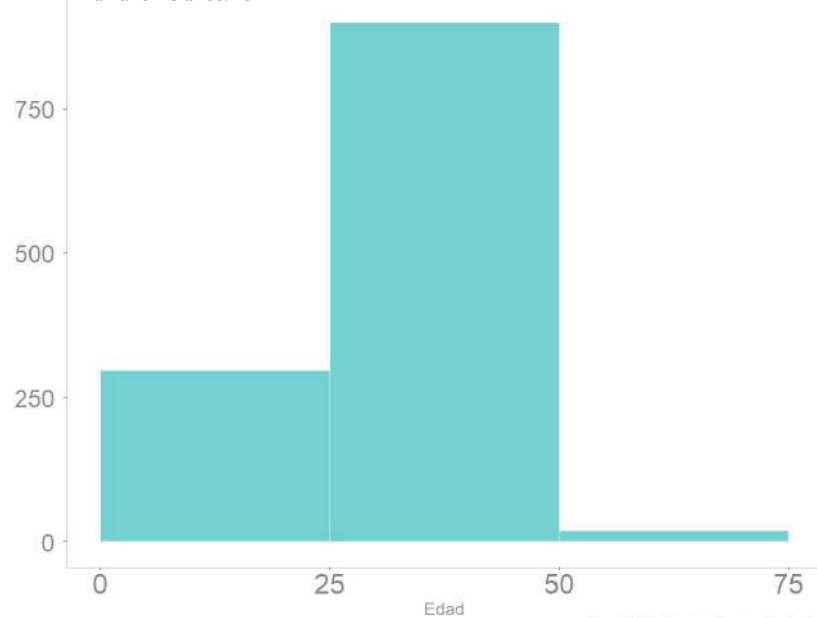
Tamaño de años: 10



Fuente: Ministerio de educación nacional de Colombia

Distribución de la edad de las personas que presentaron las pruebas saber pro (2018- abr 2022)

Tamaño de años: 25



Fuente: Ministerio de educación nacional de Colombia

Los tamaños de los contenedores de un histograma pueden modificarse.

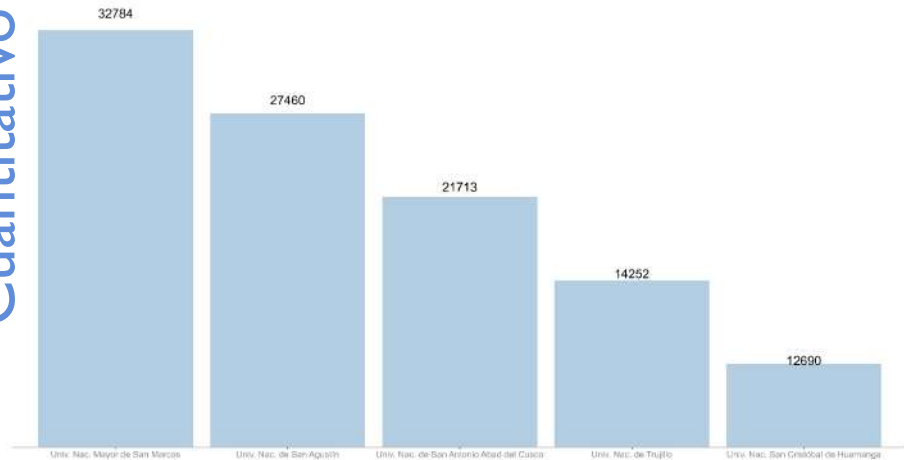


Histograma



Estudiantes matriculados en cinco universidades públicas seleccionadas de Perú 2022

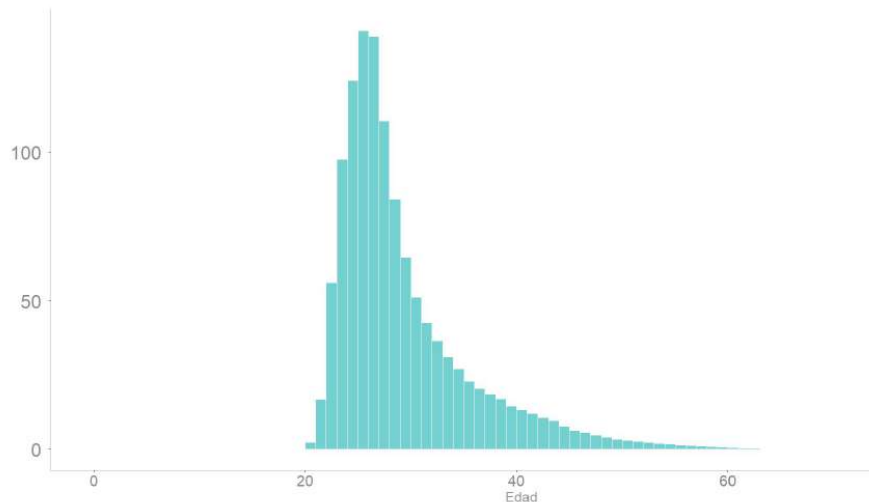
Quantitativo



Cualitativo

Fuente: INSTITUTO NACIONAL DE ESTADÍSTICA E INFORMÁTICA*

Distribución de la edad de las personas que presentaron las pruebas saber pro (2018- abr 2022)



Quantitativo

Fuente: Ministerio de Educación



El histograma, se parece mucho a un gráfico de barras, y a menudo se confunde con uno, pero son diferentes.

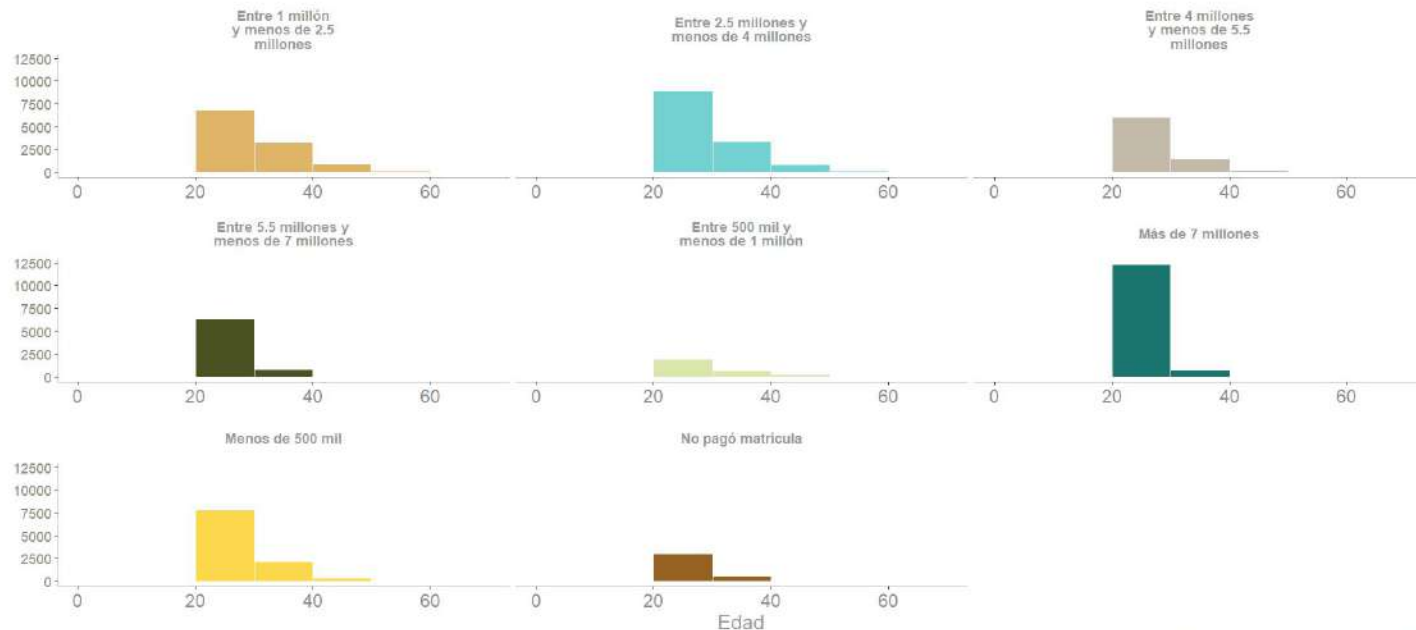


Histograma

Histograma:



Distribución de edad de las personas que presentaron las pruebas saber pro conforme al valor de matrícula (2018- 2022)



Fuente: Ministerio de educación nacional de Colombia

Los histogramas no involucran variables categóricas, sin embargo, éstas se pueden incorporar.



Diagrama de cajas

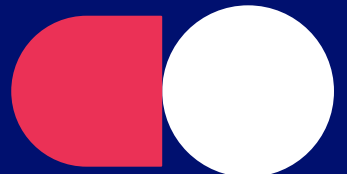




Diagrama de cajas

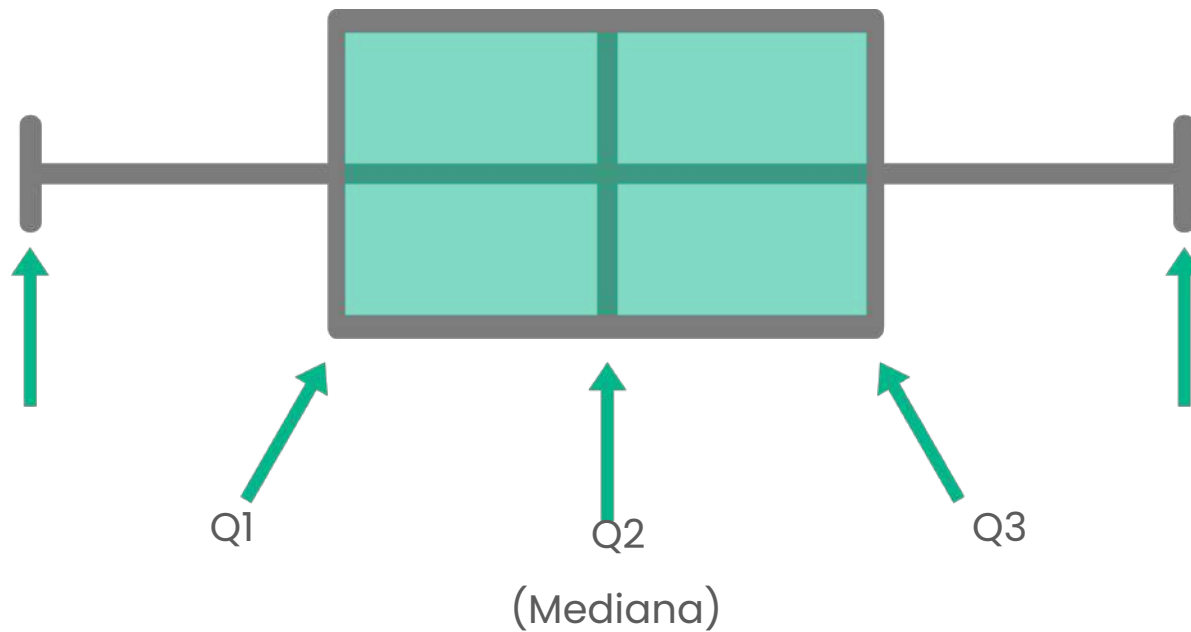




Diagrama de cajas

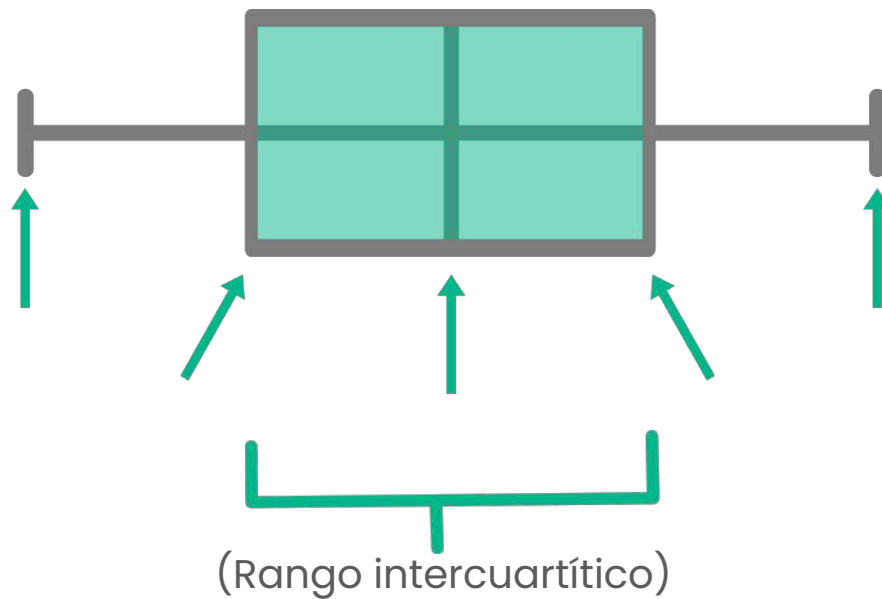




Diagrama de cajas

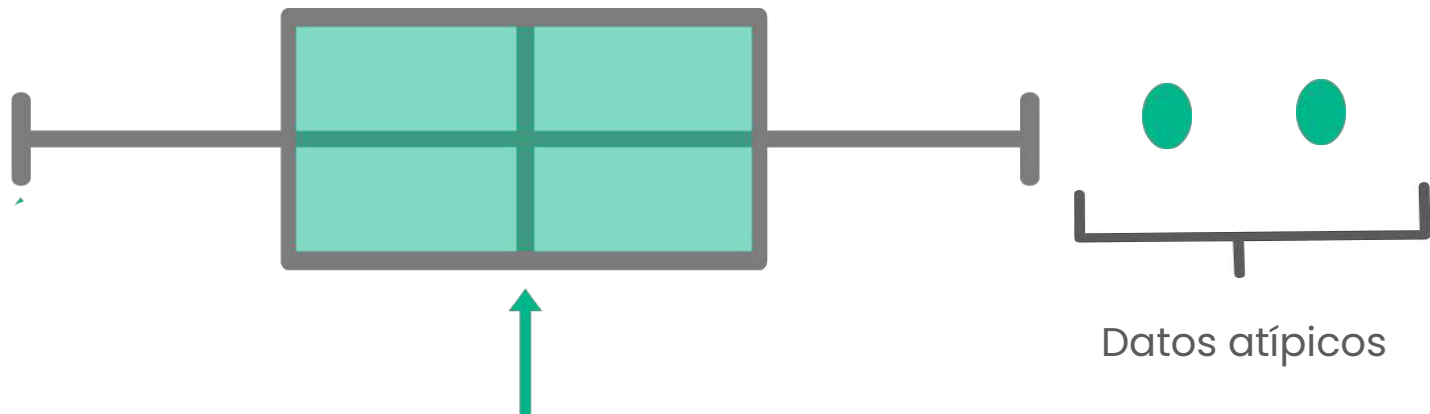
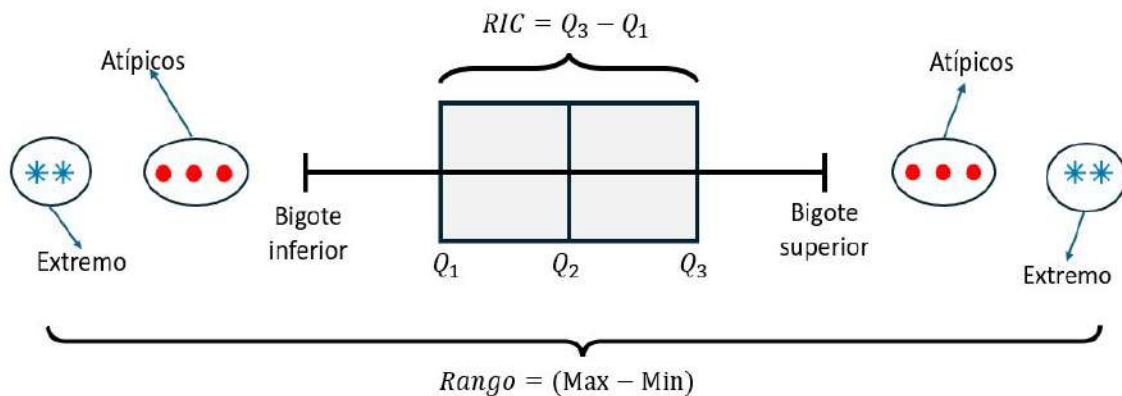




Diagrama de cajas



Atípicos:

- $L_i = Q_1 - 1,5 * RIC$
- $L_s = Q_3 + 1,5 * RIC$

Extremos:

- $L_i = Q_1 - 3 * RIC$
- $L_s = Q_3 + 3 * RIC$



Detección de valores atípicos



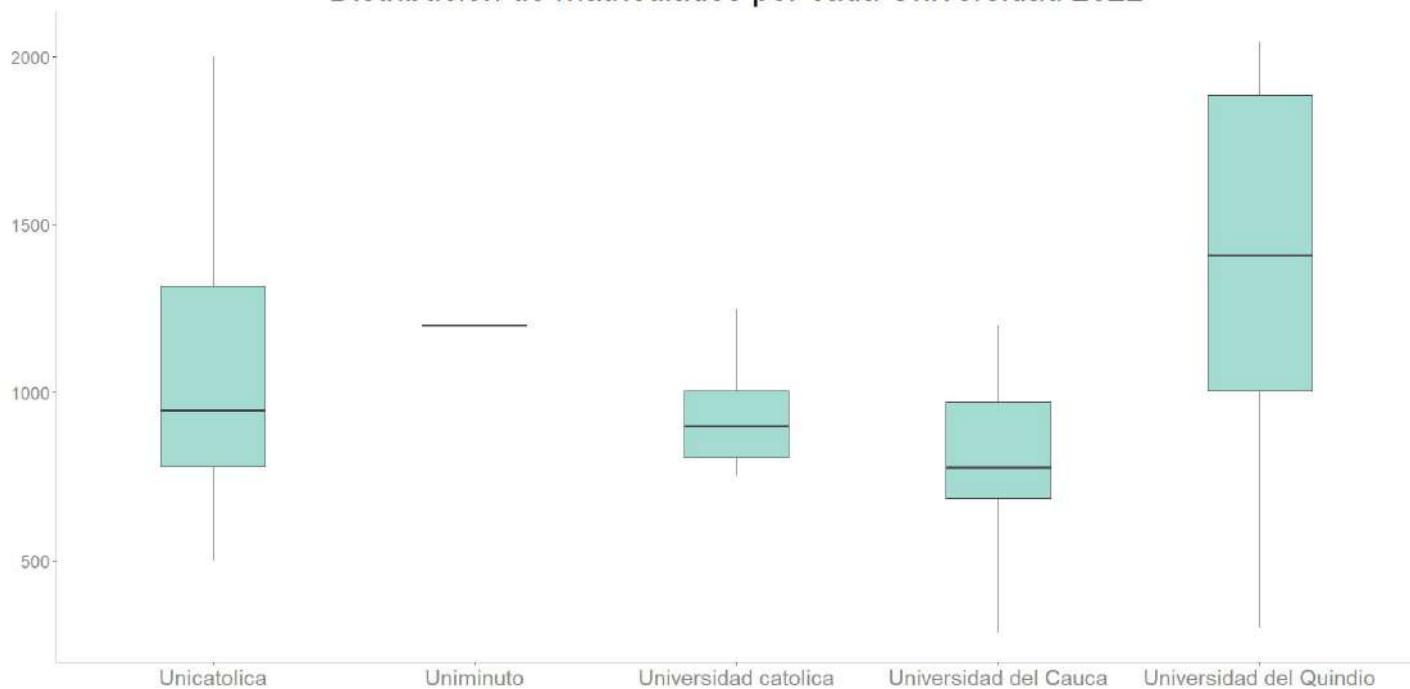
- Calcular $Q1$ y $Q3$.
- Calcular el RIC ($Q3-Q1$).
- Identificar los límites:
 - Límite inferior: $Q1 - 1.5 * RIC$
 - Límite superior: $Q3 + 1.5 * RIC$

Cualquier valor que esté por debajo del Límite Inferior o por encima del Límite Superior se considera un valor atípico.



Ejemplo

Distribución de matriculados por cada Universidad 2022



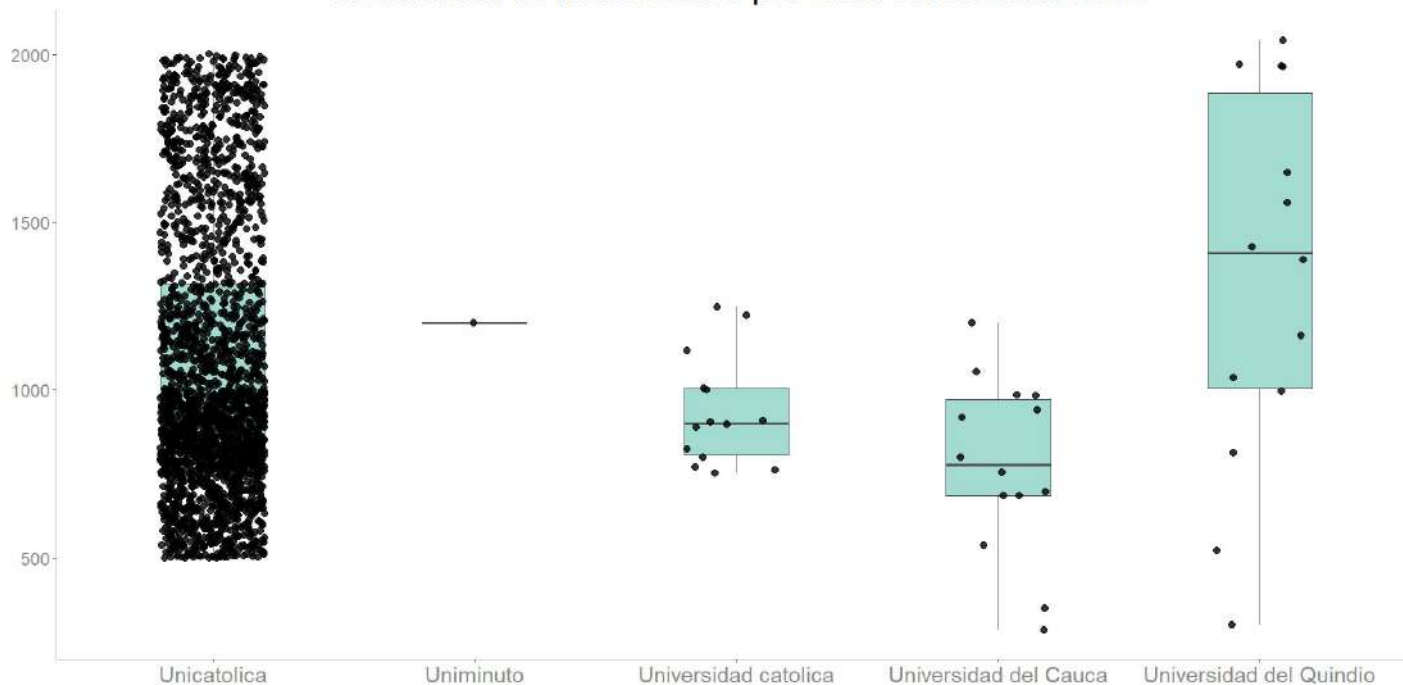
Fuente: Datos inventados*



No nos muestra la cantidad de observaciones



Distribución de matriculados por cada Universidad 2022



Fuente: Datos inventados*



Ejemplo

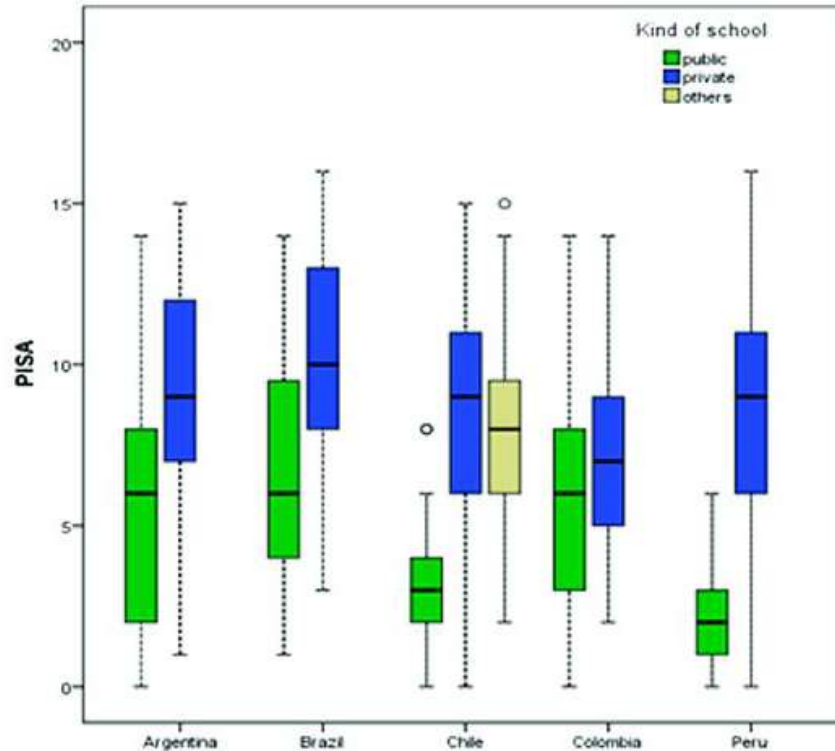




Gráfico de relación

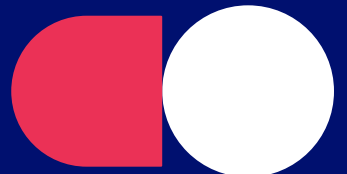




Gráfico de dispersión



Número de docentes contratados vs Número de estudiantes inscritos en el departamento Valle del Cauca 2022 "

El gráfico de dispersión nos permite visualizar patrones o tendencias en los datos y determinar si existe una relación entre las dos variables.

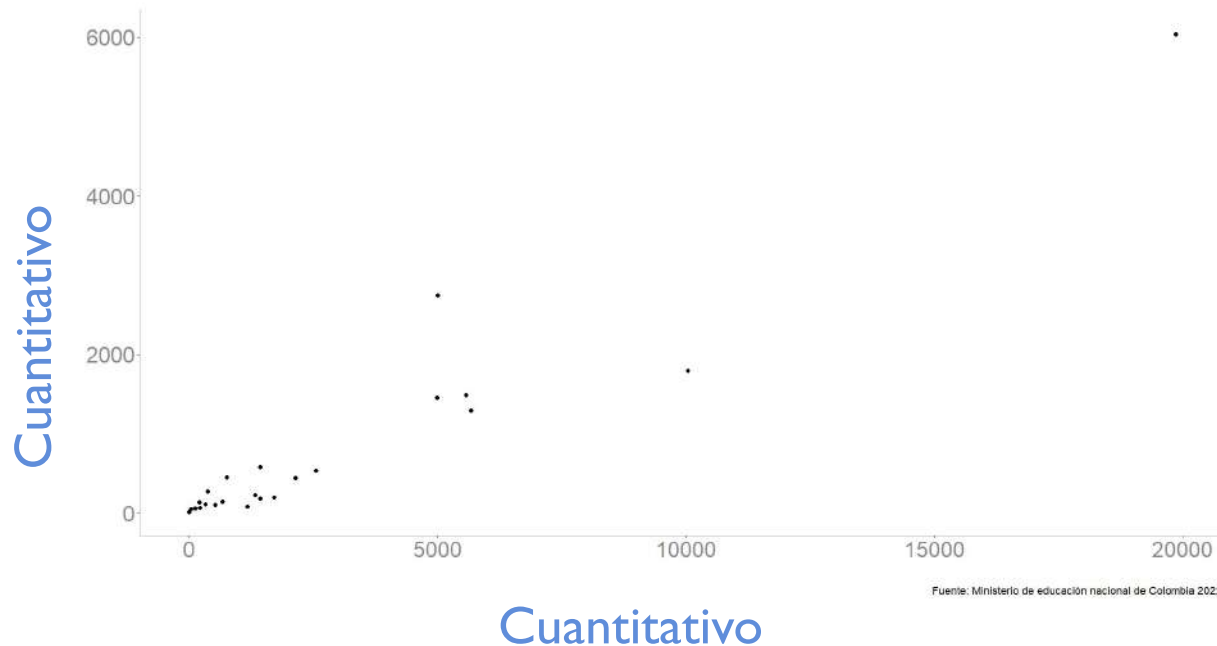
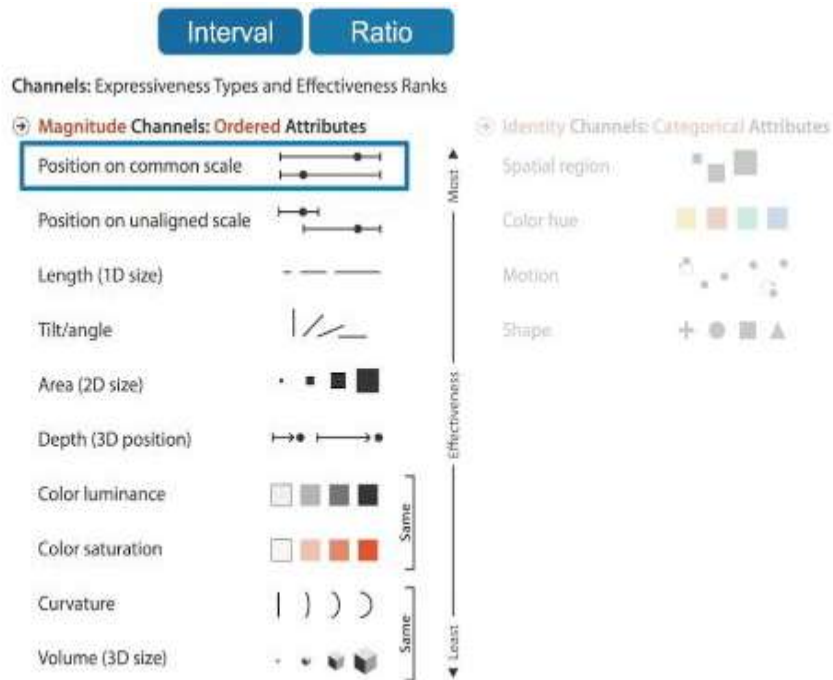




Gráfico de dispersión



Visualization Analysis and Design. Tamara Munzner, with illustrations by Eamonn Maguire. A K Peters Visualization Series, CRC Press, 2014.

De acuerdo a la codificación, el gráfico de dispersión es efectivo.



Gráfico de dispersión



No necesariamente los cambios en la variable independiente causen cambios en la variable dependiente, sino más bien se están eligiendo un valor para x como entrada y luego están viendo qué sucede con el valor de y como salida.

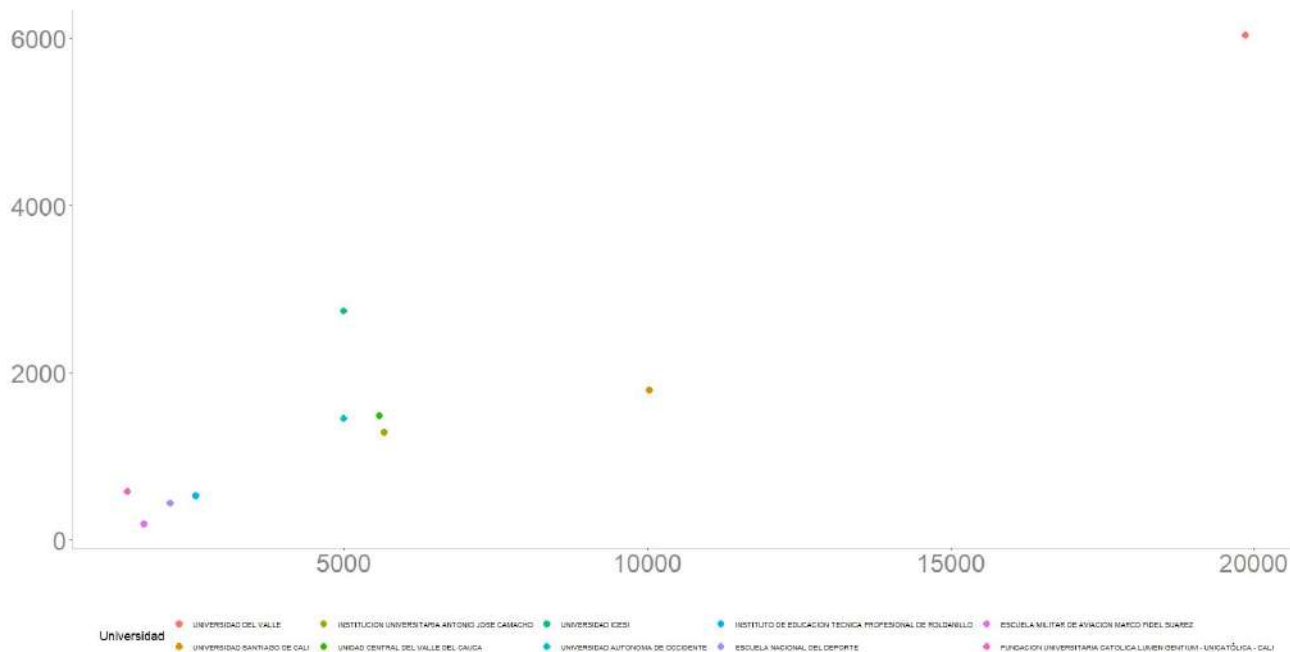
.



Gráfico de dispersión



Número de docentes contratados vs Número de estudiantes inscritos por el top 10 instituciones universitarias con más inscripciones en el departamento Valle del Cauca en 2022



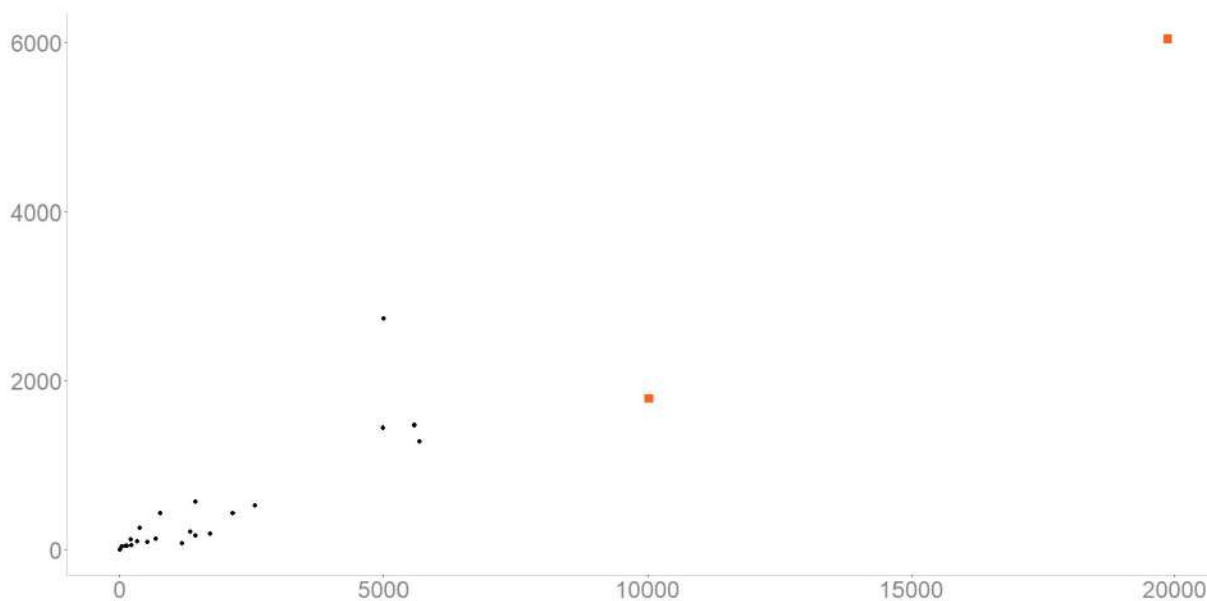
El gráfico de dispersión puede mostrar varios tipos de relaciones entre las variables.



Gráfico de dispersión



Número de docentes contratados vs Número de estudiantes inscritos en el departamento Valle del Cauca 2022 "

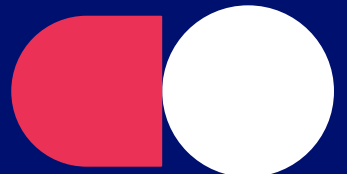


Fuente: Ministerio de educación nacional de Colombia 2022 "

El gráfico de dispersión nos ayuda a identificar valores atípicos o puntos que se alejan significativamente del patrón general de los datos.



Actividad





Actividad



Grupos: individual

Descripción: El gráfico que han traído lo van a evaluar de la siguiente forma:

Aspecto	Pregunta de Evaluación
Título	¿El gráfico tiene un título claro y específico que indica qué muestra?
Ejes	¿Los ejes están claramente etiquetados y tienen unidades (si aplica)?
Tipo de gráfico	¿El tipo de gráfico elegido es el adecuado para el tipo de datos?
Claridad de los datos	¿Se entiende claramente qué datos se están mostrando y qué significa cada elemento del gráfico?
Uso del color	¿Los colores ayudan a interpretar los datos (sin confundir)? ¿Se evita usar colores innecesarios o distractores?
Cantidad de información	¿El gráfico muestra solo lo necesario, sin sobrecargarlo con información innecesaria?
Leyenda o etiquetas	Si hay varios elementos, ¿hay una leyenda clara o etiquetas que los expliquen?
Interpretación correcta	¿El gráfico realmente ayuda a entender mejor los datos, y no los distorsiona o confunde?



Repaso



slideswith.com/PREPASOS9

Leer



Analiza



Medir



Describir





- ¿Qué gráfico es ideal cuando se tiene un gran número de variables? R/ Gráfico de barras
- ¿Un gráfico de líneas es útil para mostrar tendencias y cambios en los datos a lo largo del tiempo? R/Verdadero
- Cuando tenemos un gráfico debemos ponerle marco. R/Falso
- Los segmentos de los gráficos circulares deberían ser R/Mutuamente excluyentes Y colectivamente exhaustivos
- ¿Cuál de las siguientes afirmaciones describe mejor el propósito principal de un gráfico de líneas? R/ Representar tendencias y cambios en los datos a lo largo del tiempo.



Gracias



