

Gráficos Estadísticos y librería plotly

I. Introducción

1. Estadística Descriptiva

1.1. Variables estadísticas

1.2. Gráficos

1.2.1. Tipos de gráficos: linea, barras, circular, etc.

1.2.2. Ejes de datos

1.2.3. Tipos de gráficos de barras en detalle :sencillo,etc..

- Practica P01

II. Herramientas en Python

2. Librería Plotly

2.1. Gráficas de líneas

- Practica P02: ejercicios 1

2.2. Gráfico de Barras

- Practica P02: ejercicios 2

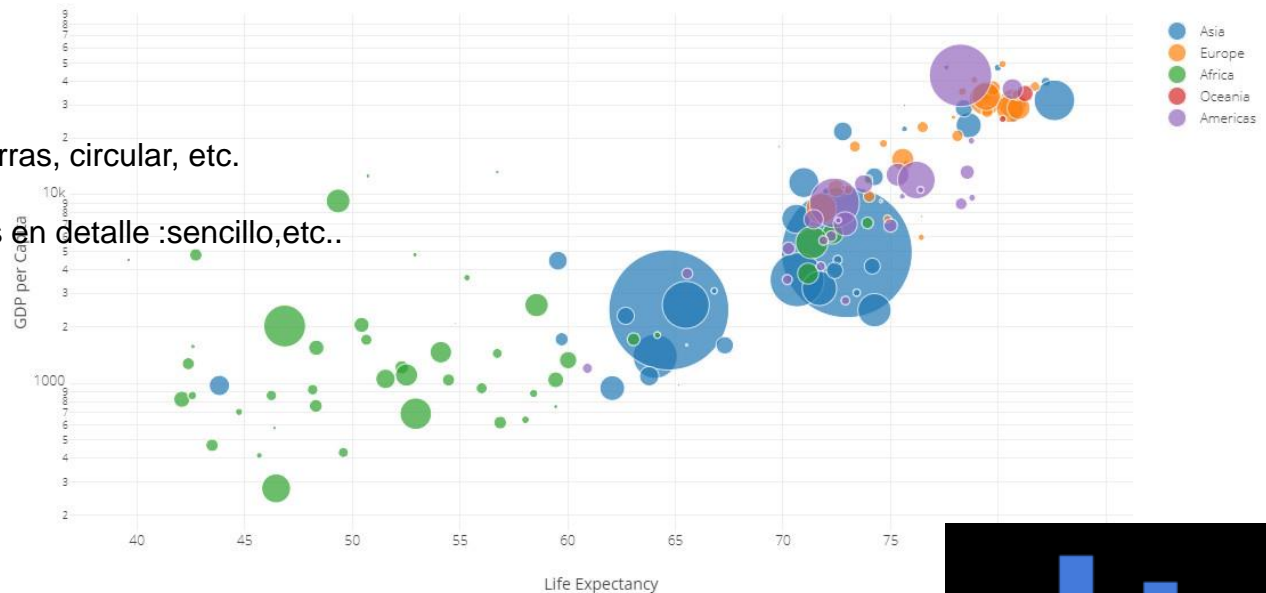
2.3. Tartas o gráficos circulares

- Practica P02: ejercicios 3

2.4. Otros gráficos básicos

2.5. Gráficos Estadísticos

- Practica p03



1. Estadística Descriptiva

Índice

- Variables estadísticas

- Tipos de Gráficos

1. - Lineal
2. - De barras
3. - Circular
4. - Polígono de frecuencias
5. - Histograma

- Cambios de ejes

- Ejemplo de histograma

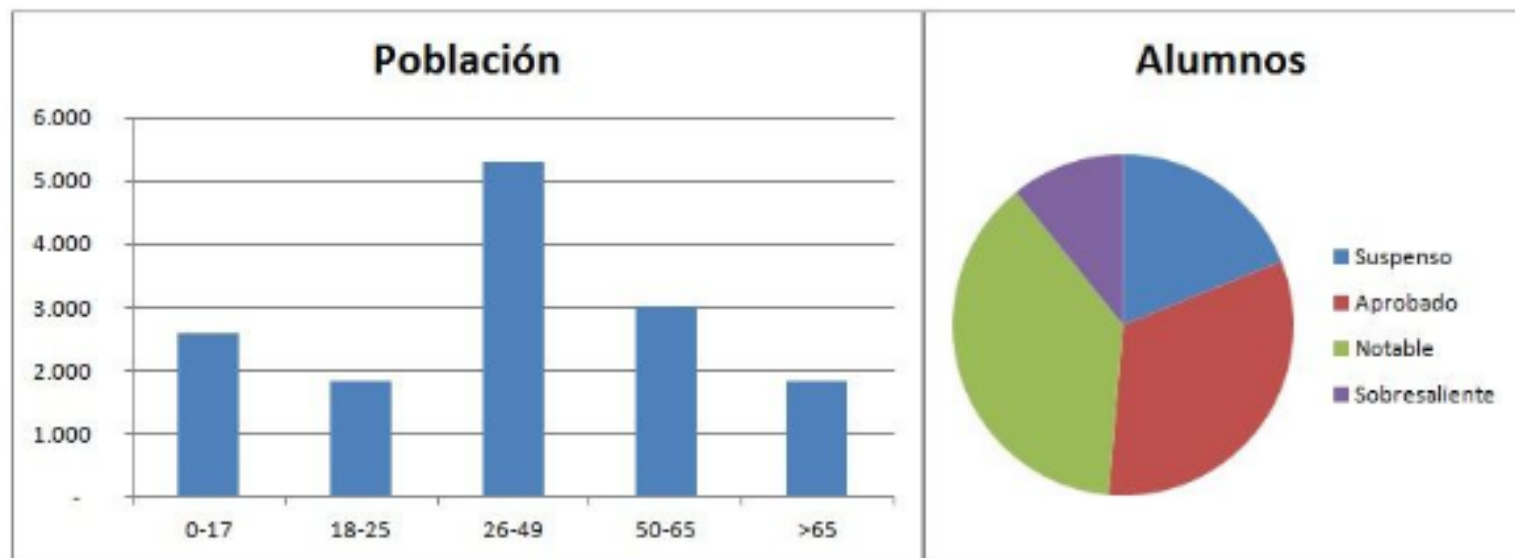


- Tipos de gráficos de Barras

- Sencillo
- Agrupado
- Apilado
- Pirámide

¿Qué es la estadística descriptiva?

La **estadística descriptiva** es la rama de la estadística que recolecta, analiza y caracteriza un conjunto de datos (peso de la **población**, beneficios diarios de una empresa, temperatura mensual,...) con el objetivo de **describir** las características y comportamientos de este conjunto mediante **medidas de resumen, tablas o gráficos**.



| Edad | Población |
|-------|-----------|
| 0-17 | 2.592 |
| 18-25 | 1.834 |
| 26-49 | 5.314 |
| 50-65 | 3.012 |
| >65 | 1.839 |

| Nota | Alumnos |
|---------------|---------|
| Suspenso | 7 |
| Aprobado | 12 |
| Notable | 14 |
| Sobresaliente | 4 |

1.1. Variables estadísticas

Una **variable estadística** es el conjunto de valores que puede tomar cierta característica de la **población** sobre la que se realiza el estudio estadístico y sobre la que es posible su medición. Estas variables pueden ser: la edad, el peso, las notas de un examen, los ingresos mensuales, las horas de sueño de un paciente en una semana, el precio medio del alquiler en las viviendas de un barrio de una ciudad, etc.

Las **variables estadísticas** se pueden clasificar por diferentes criterios. Según su medición existen dos tipos de variables:

- **Cualitativa** (o categórica): son las variables que pueden tomar como valores cualidades o categorías.

Ejemplos: ■ *Sexo (hombre, mujer)* ■ *Salud (buena, regular, mala)*

- **Cuantitativas** (o numérica): variables que toman valores numéricos.

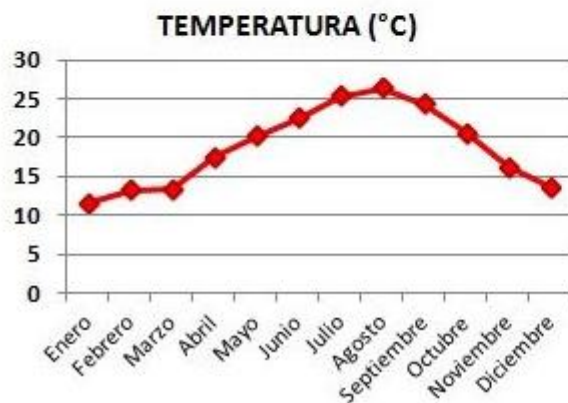
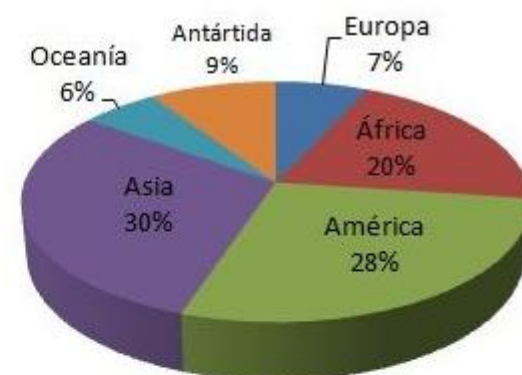
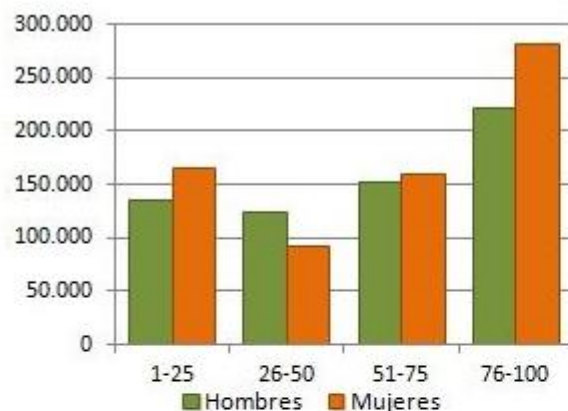
Ejemplos: ■ *Número de casas (1, 2,...). Discreta.*

■ *Temperatura : 20,5 ; 12,7 ; 30,2*

1.2. Gráficos

Un **gráfico** (o **gráfica**) es el recurso de representar los datos numéricos por medio de líneas, diagramas, dibujos, etc.

La representación gráfica es un importante suplemento al análisis y estudio estadístico.



1.2.1. Tipos de gráficos

Gráfico lineal

El **gráfico lineal** (gráfico de líneas o **diagrama lineal**) se compone de una serie de datos representados por puntos, unidos por segmentos lineales. Mediante este gráfico se puede comprobar rápidamente el cambio de tendencia de los datos.

El **diagrama lineal** se suele utilizar con **variables cuantitativas**, para ver su comportamiento en el transcurso del tiempo. Por ejemplo, en las **series temporales** mensuales, anuales, trimestrales, etc.

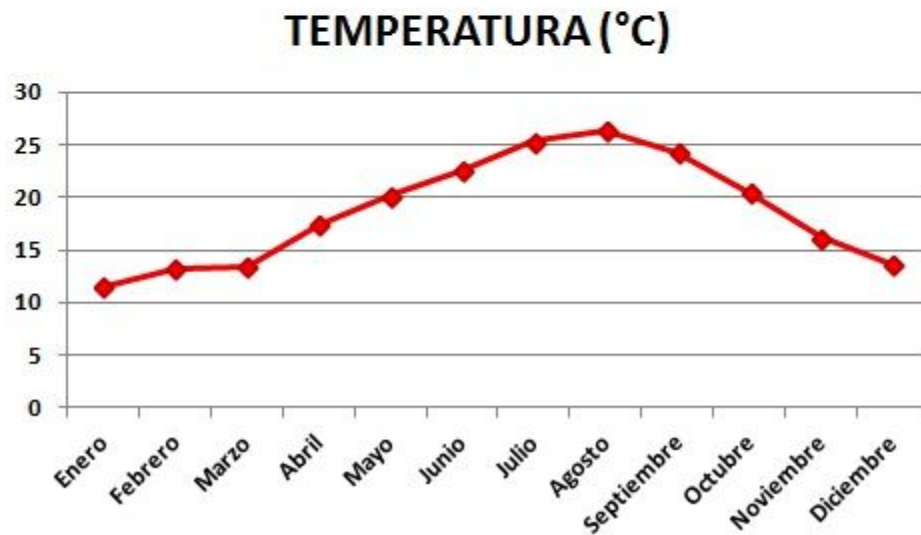


Diagrama de barras

El **diagrama de barras** es un gráfico que se utiliza para representar datos de **variables cualitativas** o discretas. Está formado por **barras** rectangulares cuya altura es proporcional a la frecuencia de cada uno de los valores de la variable.



- Las barras pueden ser **horizontales o verticales**, según si los valores de la **variable** se reflejan en el eje horizontal o vertical.
- Todas las barras deben tener el **mismo ancho** y no deben superponerse las unas con las otras.



Diagrama de barras verticales

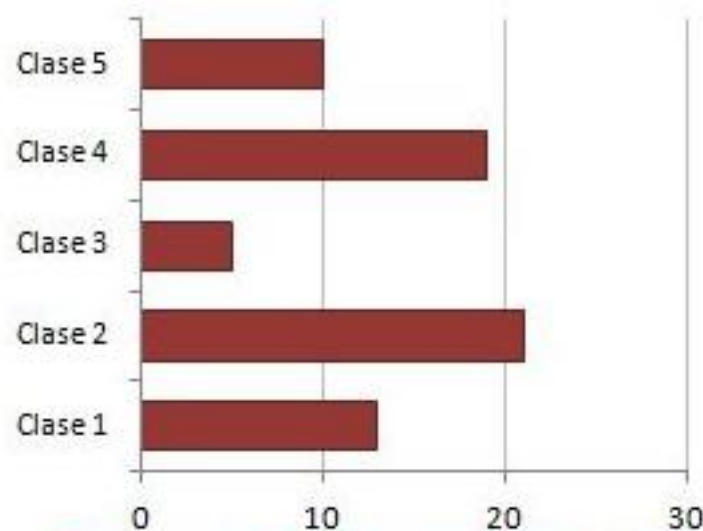
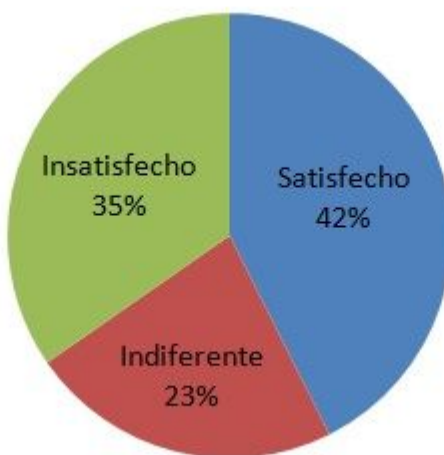


Diagrama de barras horizontales

Diagrama circular

El **diagrama circular** (también llamado **diagrama de sectores** o **diagrama de pastel**) sirve para representar **variables cualitativas** o discretas. Se utiliza para representar la proporción de elementos de cada uno de los valores de la variable.

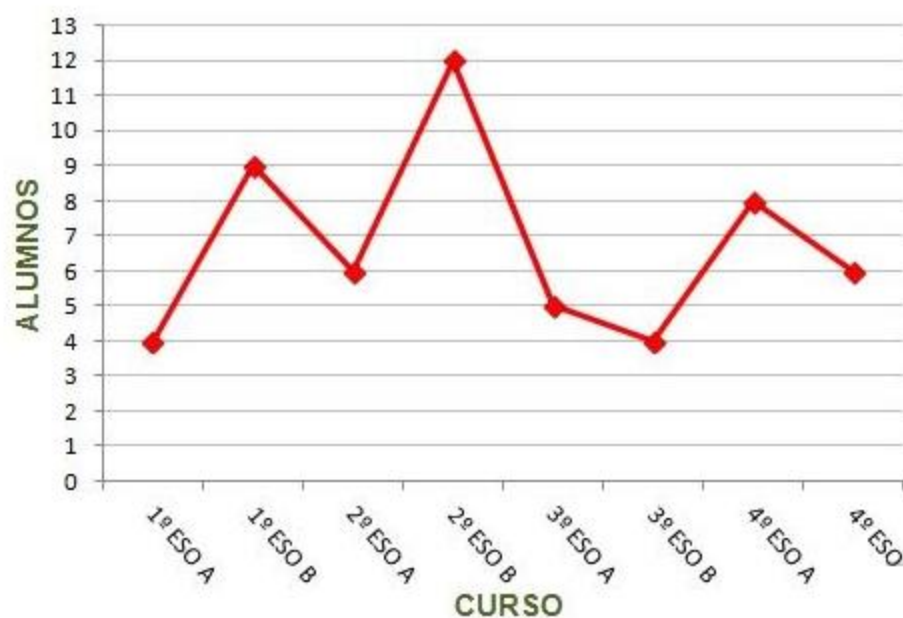
Consiste en partir el **círculo** en porciones proporcionales a la **frecuencia relativa**. Entiéndase como porción la parte del **círculo** que representa a cada valor que toma la variable.



Polígono de frecuencias

El **polígono de frecuencias** es un gráfico que permite la rápida visualización de las frecuencias de cada una de las categorías del estudio.

Normalmente se utiliza el **polígono de frecuencias** con **frecuencias absolutas**, pero también se utiliza con **frecuencias relativas**.

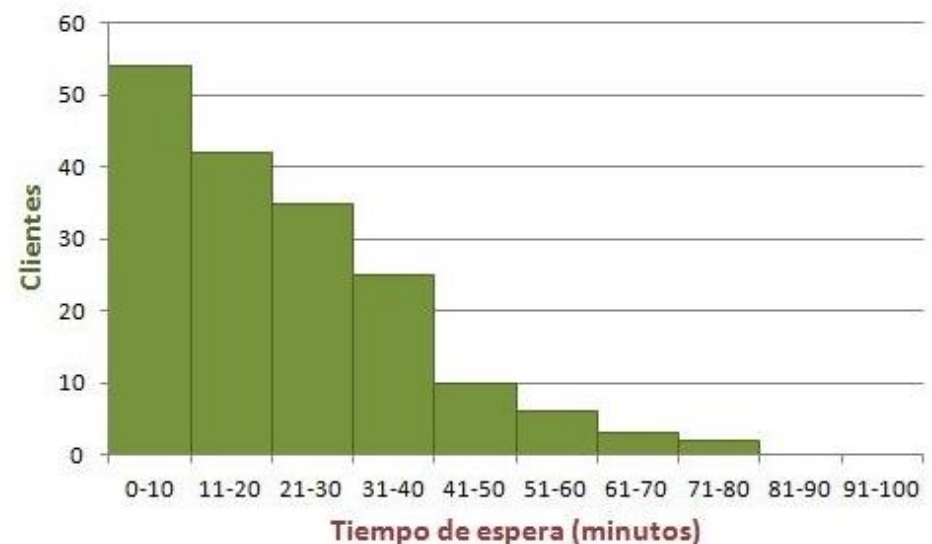


Histograma

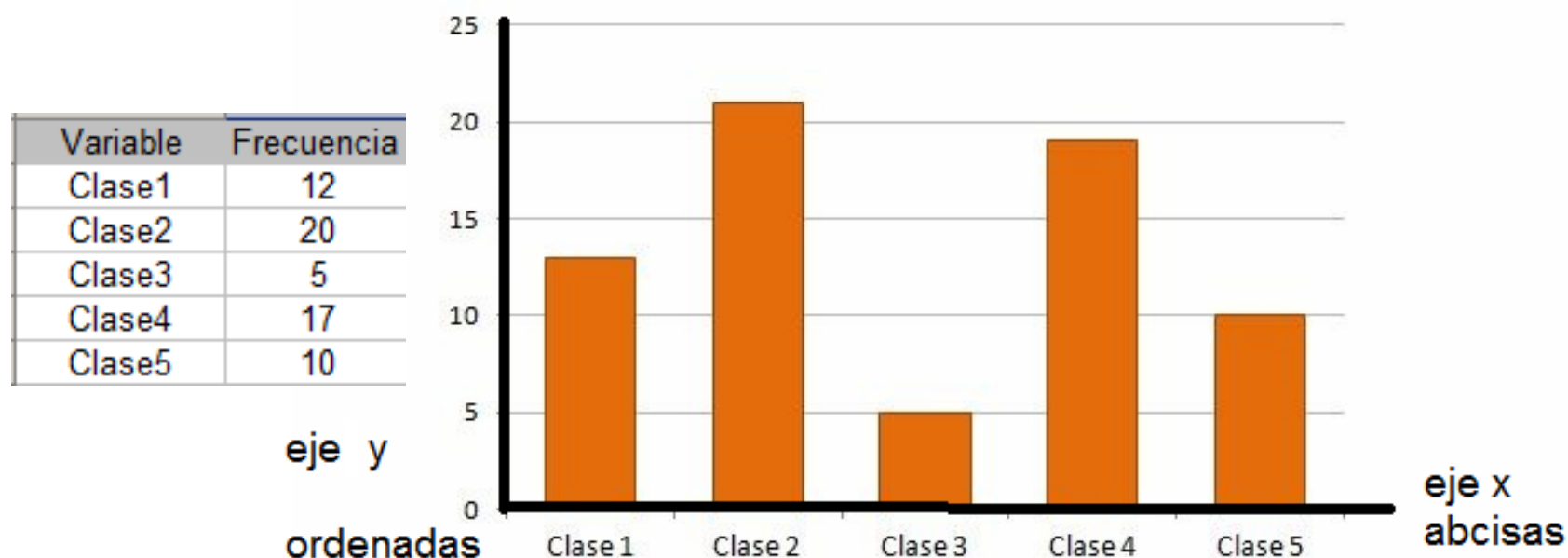
Un **histograma** es una representación gráfica de **datos agrupados** mediante intervalos. Los datos provienen de una **variables cuantitativas** continuas. Gracias a él puedes hacerte rápidamente una idea de la distribución de los datos o **muestra**.

También cabe emplear **variables cualitativas** ordinales, siendo necesario que el número de datos sea alto.

Un **histograma** es un conjunto de **rectángulos** que representan las **frecuencias absolutas** de cada uno de los intervalos. Los **intervalos** abarcan todo el conjunto sin cortarse, de manera que un elemento está solo en un intervalo.



1.2.2. Ejes de datos



Las principales características del diagrama de barras son:

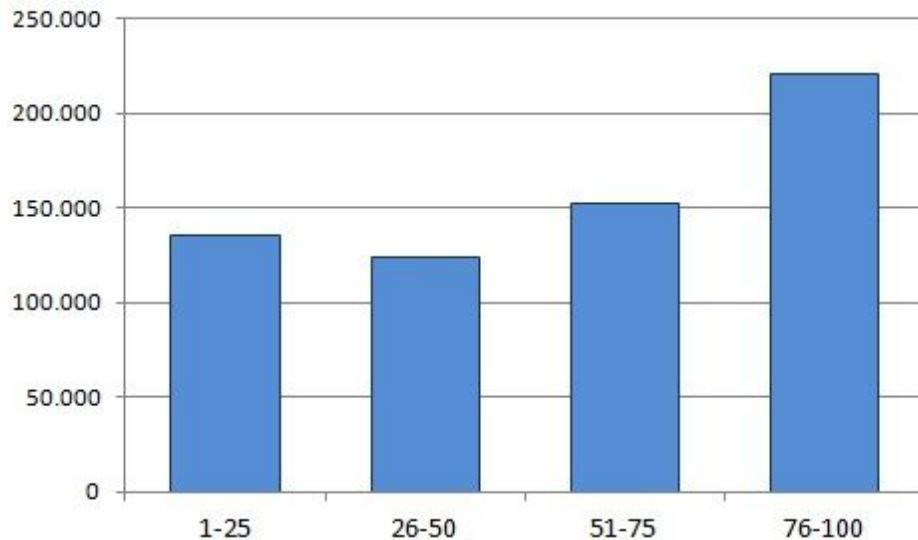
- En el **eje de abcisas** se colocan las cualidades de la variable, si la variable es cualitativa, o los valores de dicha variable, si es discreta.
- En el **eje de ordenadas** se colocan las barras proporcionales a la frecuencia relativa o absoluta del dato.

1.2.3. Tipos de gráficos de barras en detalle

- Existen cuatro tipos de **gráficos de barras** según las series de datos y como están estas representadas:

1. Gráfico de barras sencillo

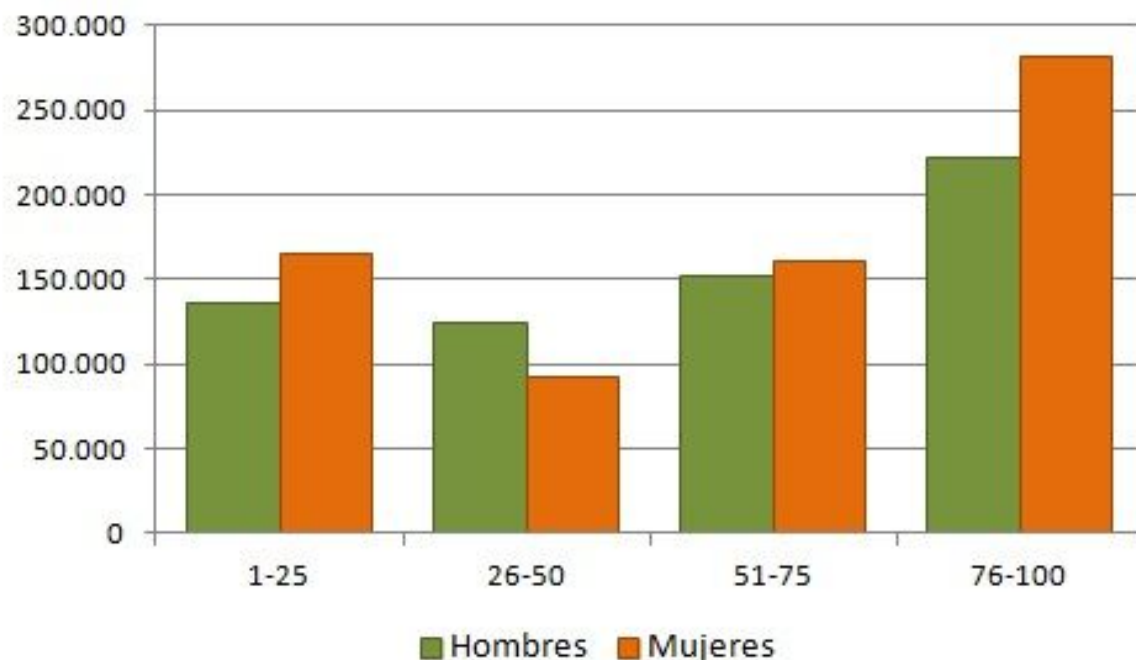
Representa los datos de **una única serie o conjunto de datos**.



- Sencillo
- Agrupado
- Apilado
- Pirámide

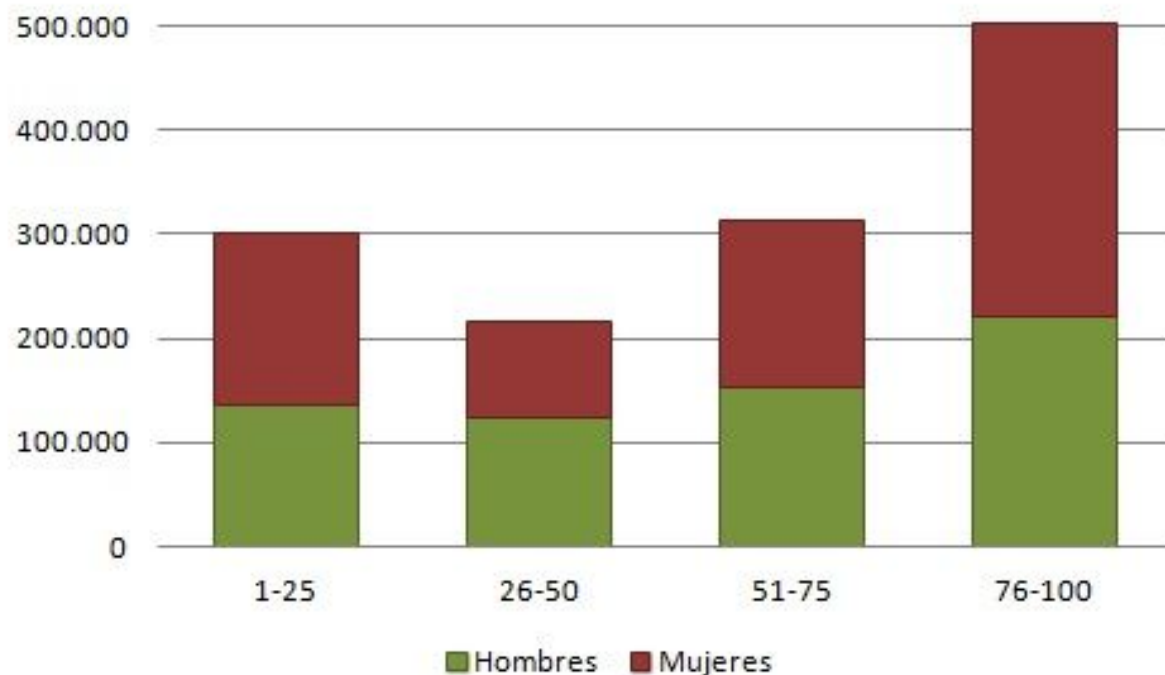
2. Gráfico de barras agrupado

- Representa los datos de **dos o más** series o **conjuntos de datos**.
- Cada **serie** se representa en un **mismo color**.
- Las barras se colocan una al lado de la otra por categoría de la variable para comparar las series de datos.



3. Gráfico de barras apilado

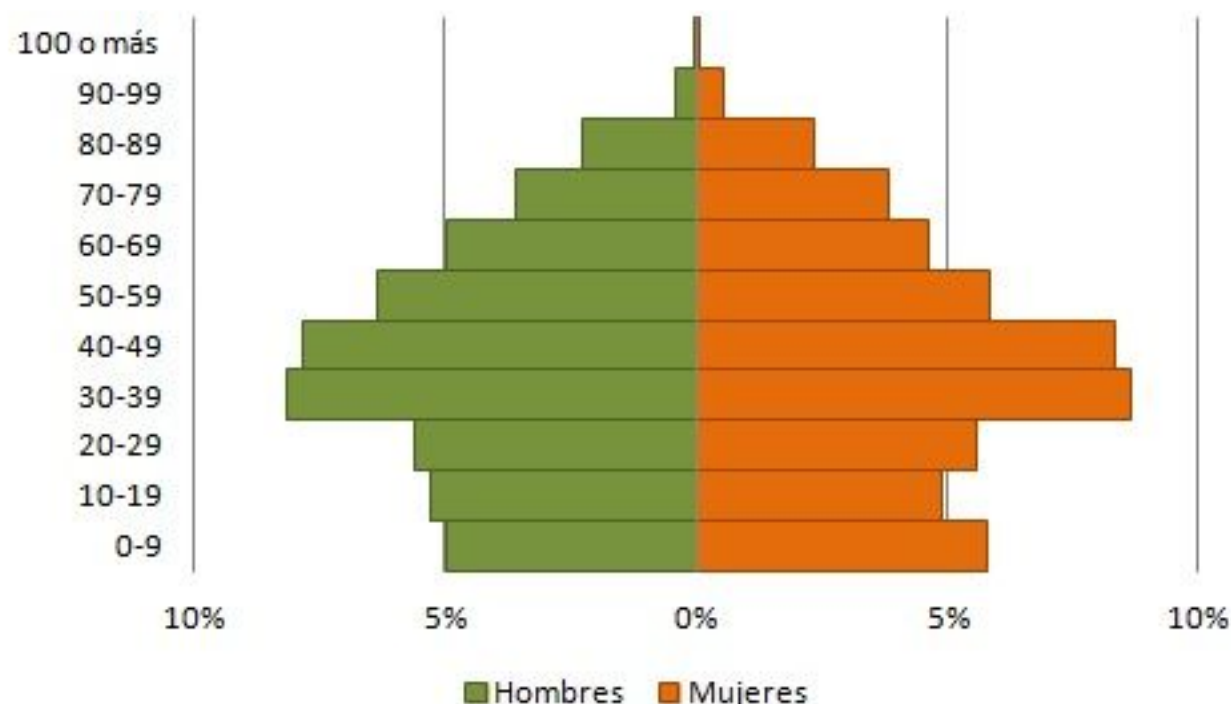
- Representa los datos de **dos o más** series o **conjuntos de datos**.
- Cada **serie** se representa en un **mismo color**.
- Cada barra representa una categoría de la variable, y se divide en segmentos que representan cada una de las series de datos.



4. Pirámide de población

Una **pirámide de población** es un **diagrama de barras bidireccional** que muestra la población por sexo y rangos de edad en un momento determinado.

En el eje vertical se representan los intervalos de edades. En el eje horizontal se representan los porcentajes de población. En una dirección se representarán los hombres y en el otro las mujeres.



Practica P01

1. Esquematiza (sobre papel) los 4 tipos de gráficos de barras usando estos datos del Instituto nacional de estadística.

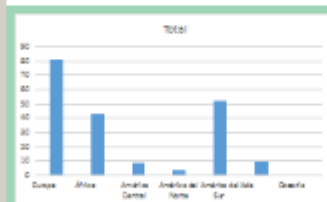
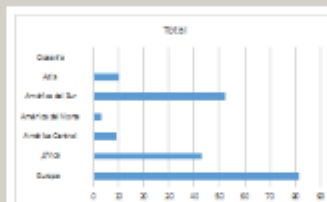
| Zona | Hombres | Mujeres | Total |
|-------------------|---------|---------|-------|
| Europa | 39,57 | 41,9 | 81,47 |
| África | 28,11 | 14,89 | 43 |
| América Central | 3 | 6,22 | 9,22 |
| América del Norte | 1,52 | 1,85 | 3,37 |
| América del Sur | 22,22 | 30,23 | 52,45 |
| Asia | 5,47 | 4,78 | 10,25 |
| Oceanía | 0,11 | 0,12 | 0,23 |

Cuadro 1.2. Distribución porcentual de la población extranjera residente en España por zonas según sexo. 1 de enero de 2001

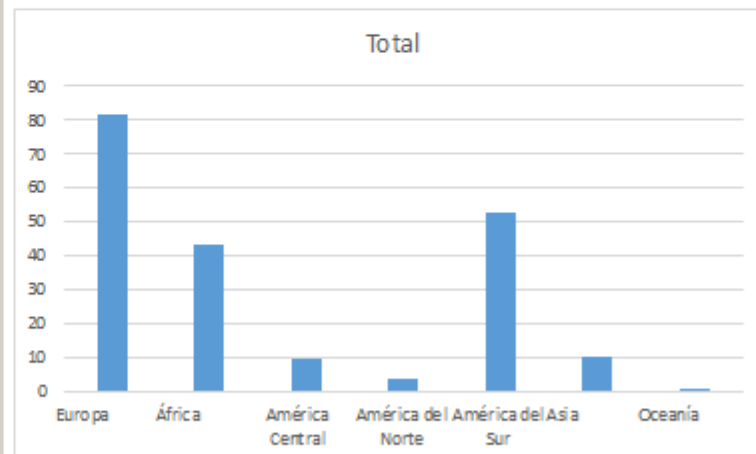
| | A | B | C | D |
|----|-------------------|---------|---------|-----|
| 1 | Zona | Hombres | Mujeres | Tot |
| 2 | Europa | 39,57 | 41,9 | 81 |
| 3 | África | 28,11 | 14,89 | |
| 4 | América Central | 3 | 6,22 | 9 |
| 5 | América del Norte | 1,52 | 1,85 | 3 |
| 6 | América del Sur | 22,22 | 30,23 | 52 |
| 7 | Asia | 5,47 | 4,78 | 10 |
| 8 | Oceanía | 0,11 | 0,12 | 0 |
| 9 | | | | |
| 10 | | | | |
| 11 | | | | |

Insertar gráfico

Gráficos recomendados Todos los gráficos

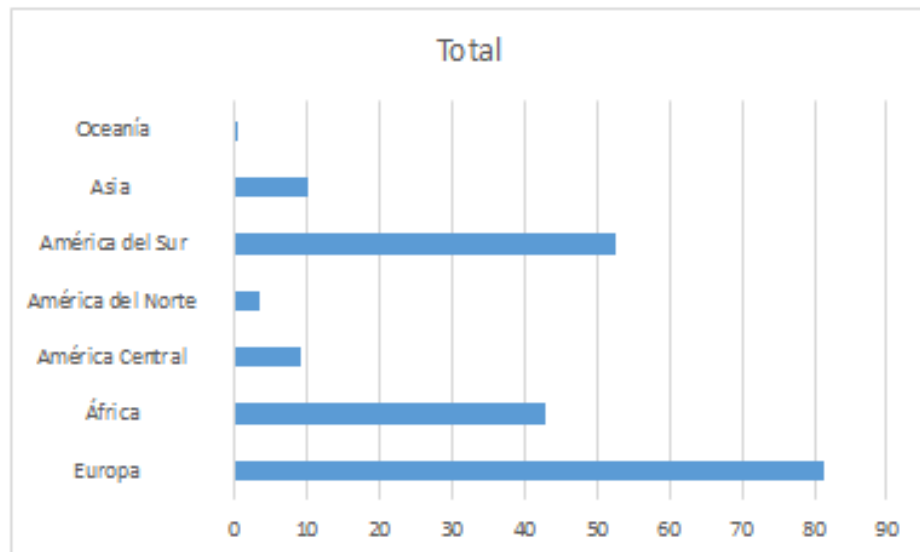


Columna agrupada



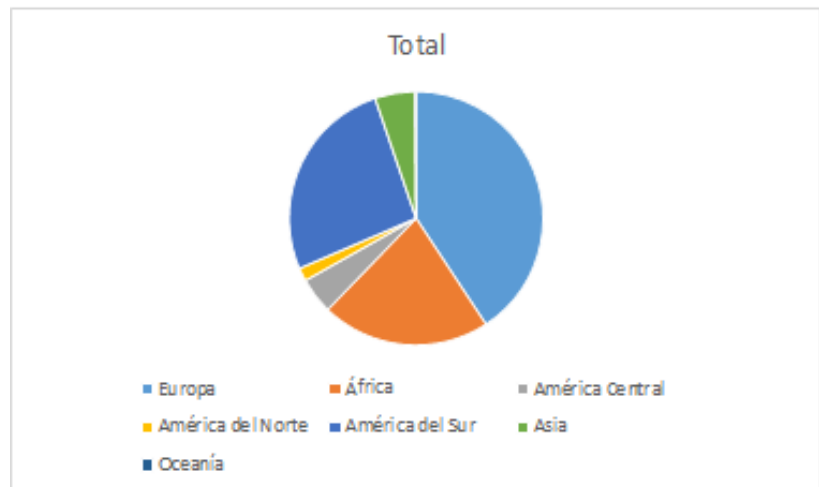
El gráfico de columnas agrupadas se usa para comparar valores entre algunas categorías. Úsalo cuando el orden de las categorías no sea importante.

Barra agrupada



El gráfico de barras agrupadas se usa para comparar valores entre algunas categorías. Úselo cuando el gráfico muestra una duración o cuando el texto de la categoría es largo.

Circular



El gráfico circular se usa para mostrar las proporciones del total. Úselo para mostrar los números que se relacionan a una suma mayor y cuando el total de los número es siempre igual a 100%. No use este gráfico si contiene muchos sectores, ya que resultaría difícil estimar los ángulos.

Insertar gráfico



Gráficos recomendados

Todos los gráficos

Reciente

Plantillas

Columna

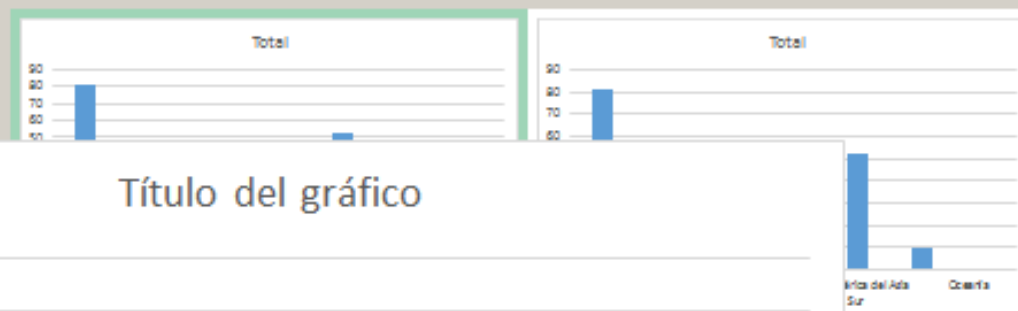
Línea

Circular

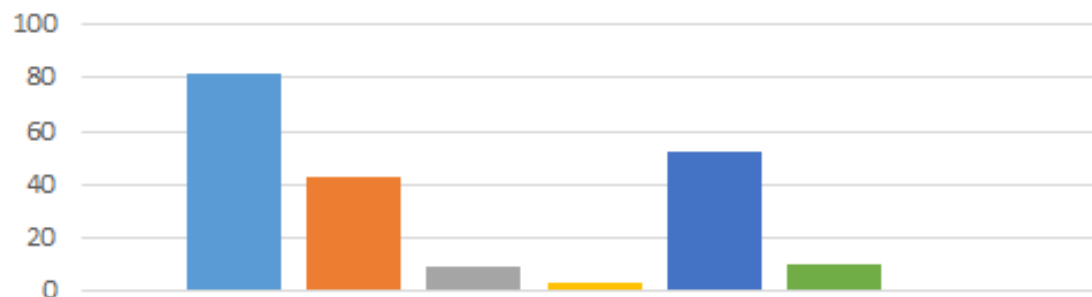
Barra



Columna agrupada



Título del gráfico











Total

| | | | | | |
|-------------------|-------|-------|---------------------|-------|-------|
| ■ Europa | 39,57 | 41,9 | ■ África | 28,11 | 14,89 |
| ■ América Central | 3 | 6,22 | ■ América del Norte | 1,52 | 1,85 |
| ■ América del Sur | 22,22 | 30,23 | ■ Asia | 5,47 | 4,78 |
| ■ Oceanía | 0,11 | 0,12 | | | |

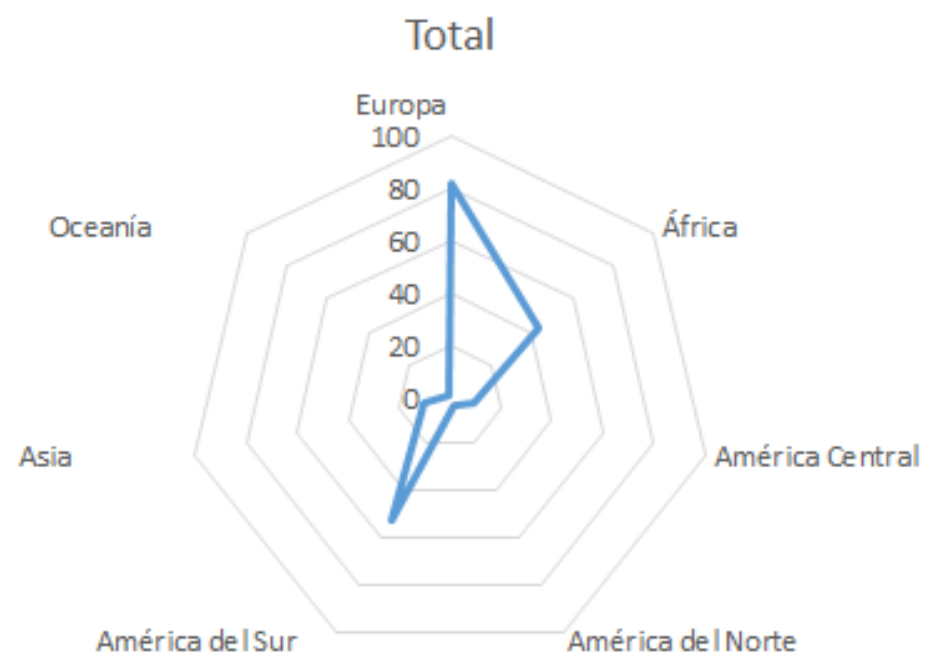
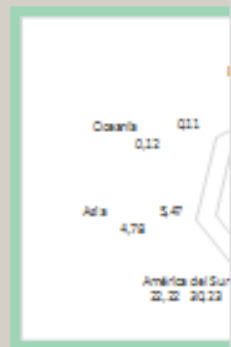
Gráficos recomendados

Todos los gráficos

-  Reciente
-  Plantillas
-  Columna
-  Línea
-  Circular
-  Barra
-  Área
-  X Y (Dispersión)
-  Cotizaciones
-  Superficie
-  Radial
-  Cuadro combinado



Radial

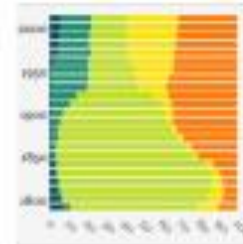




Dash -
Interactive
Python Apps



Static Image
Export



Updating
Plotly Graphs

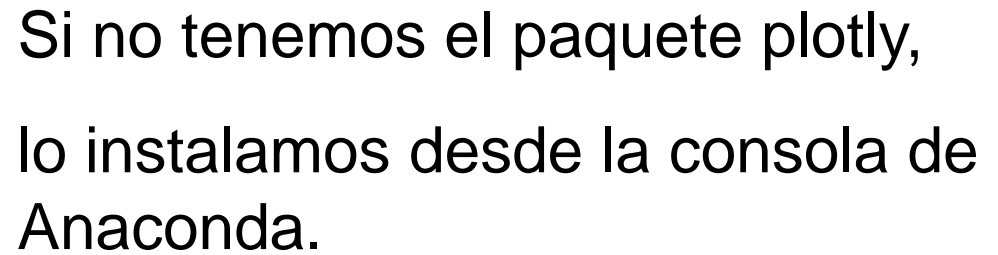


Jupyter
Notebook
Tutorial

Herramientas para estadística en Python

2. Librería Plotly : Librería gráfica

<https://plotly.com/python/basic-charts/>



```
pip install plotly==4.14.3
```

```
C:\Windows\system32\cmd.exe
Collecting plotly
  Could not find a version that satisfies the requirement plotly (from v
)
No matching distribution found for plotly

(base) C:\Users\Usuario>pip install plotly==4.14.3
Collecting plotly==4.14.3
  Downloading https://files.pythonhosted.org/packages/1f/f6/bd3c17c8003b
28e80e1acac97ed8402635e46c2571f8e1ef63af/plotly-4.14.3-py2.py3-none-any.
2MB)
100% |#####| 13.2MB 706kB/s
```

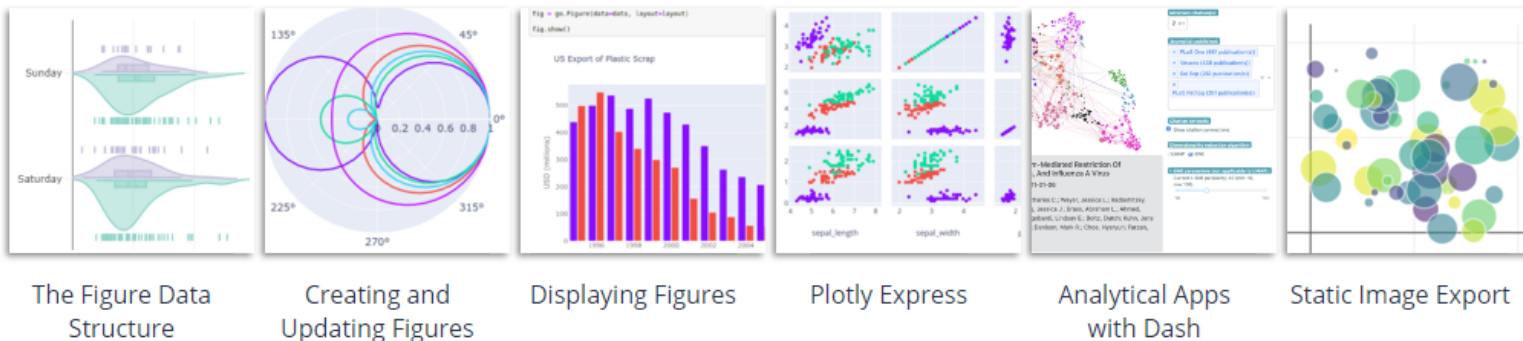
El **Python Package Index** o **PyPI** es el [repositorio](#) de software oficial para aplicaciones de terceros en el lenguaje de programación [Python](#). Los desarrolladores de Python pretenden que sea un catálogo exhaustivo de todos los [paquetes](#) de Python escritos en [código abierto](#)¹

2.1. Fundamentos

<https://plotly.com/python/>

Plotly Python Open Source Graphing Library Fundamentals

Plotly's Python graphing library makes interactive, publication-quality graphs online. Tutorials and tips about fundamental features of Plotly's python API.



La librería Plotly permite realizar todo tipo de gráficos con facilidad.

Nos centraremos en el módulo de **objetos gráficos** y la clase **Figura**.

```
import plotly.graph_objects as go
```

[cursory/plotly.ipynb](https://github.com/plotly/python-api/blob/master/plotly/figure_factory/figure.py)

```
fig = go.Figure()
```


Dash

GitHub

community.plotly.com

Examples

Fundamentals

Basic Charts

Statistics

Artificial

Machine

Science

Finance

Maps

3D Charts

Subplots

Animation

Charts

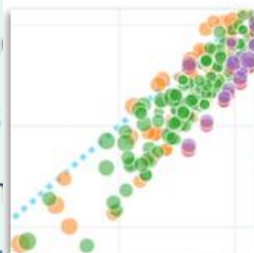
Advanced

Sobre Figure, exploramos los gráficos Básicos

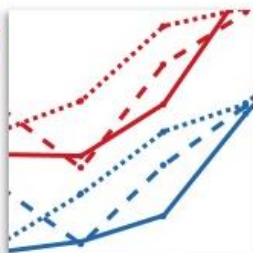
- de líneas : `go.Scatter (x[], y[], name, layout{})`
- de barras : `go.Bar (x[], y[], name, layout{})`
- y tartas : `go.Pie (labels, values, ..., ...)`

Plotly Python Open Source Graphing Library Basic Charts

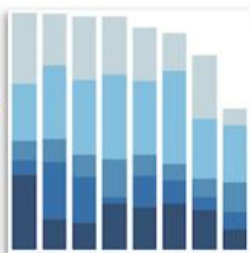
Plotly's Python graphing library makes interactive, publication-quality graphs online. Examples of how to make basic charts.



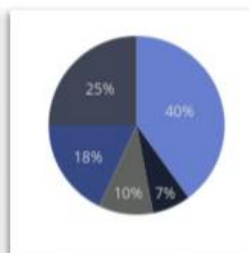
Scatter Plots



Line Charts



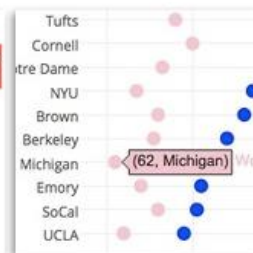
Bar Charts



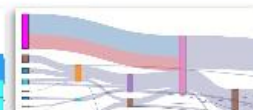
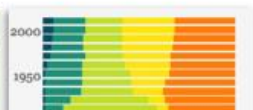
Pie Charts



Bubble Charts



Dot Plots



2.1. Gráficas de líneas

Practica P02

Ejercicio 1. Gráfico de líneas.

Ejecuta este programa en tu cuaderno y después de comprobar el total elimina la línea de `print` y añade el código de la siguiente diapositiva para mostrar una gráfica de líneas.

```
import plotly.graph_objects as go

#import plotly.io as pio
#pio.renderers.default = "browser" ←----- Importante para jupyter
# ----- Datos
zona = ['Europa', 'África', 'América Nort', 'América Cent',
        'América Sur', 'Asia', 'Oceanía']

hombres = [39.57, 28.11, 3, 1.51, 22.22, 5.47, 0.11]
mujeres = [41.9, 14.89, 6.22, 1.85, 30.23, 4.78, 0.12]
total = [round(x + y, 2) for x, y in zip(hombres, mujeres)]

print(total)
```

```

fig = go.Figure()

# Create and style traces
# dash options include 'dash', 'dot', and 'dashdot'

fig.add_trace(go.Scatter(x=zona, y=hombres, name='Hombres',
                        line=dict(color='firebrick', width=2,
dash='dot'))))
fig.add_trace(go.Scatter(x=zona, y=mujeres, name = 'Mujeres',
                        line=dict(color='royalblue', width=2)))
fig.add_trace(go.Scatter(x=zona, y=total, name='Total',
                        line=dict(color='firebrick', width=2,
dash='dash'))))

# Edit the layout
fig.update_layout(title='Residentes por Zona origen datos 2001',
                  xaxis_title='Zona',
                  yaxis_title='Personas')

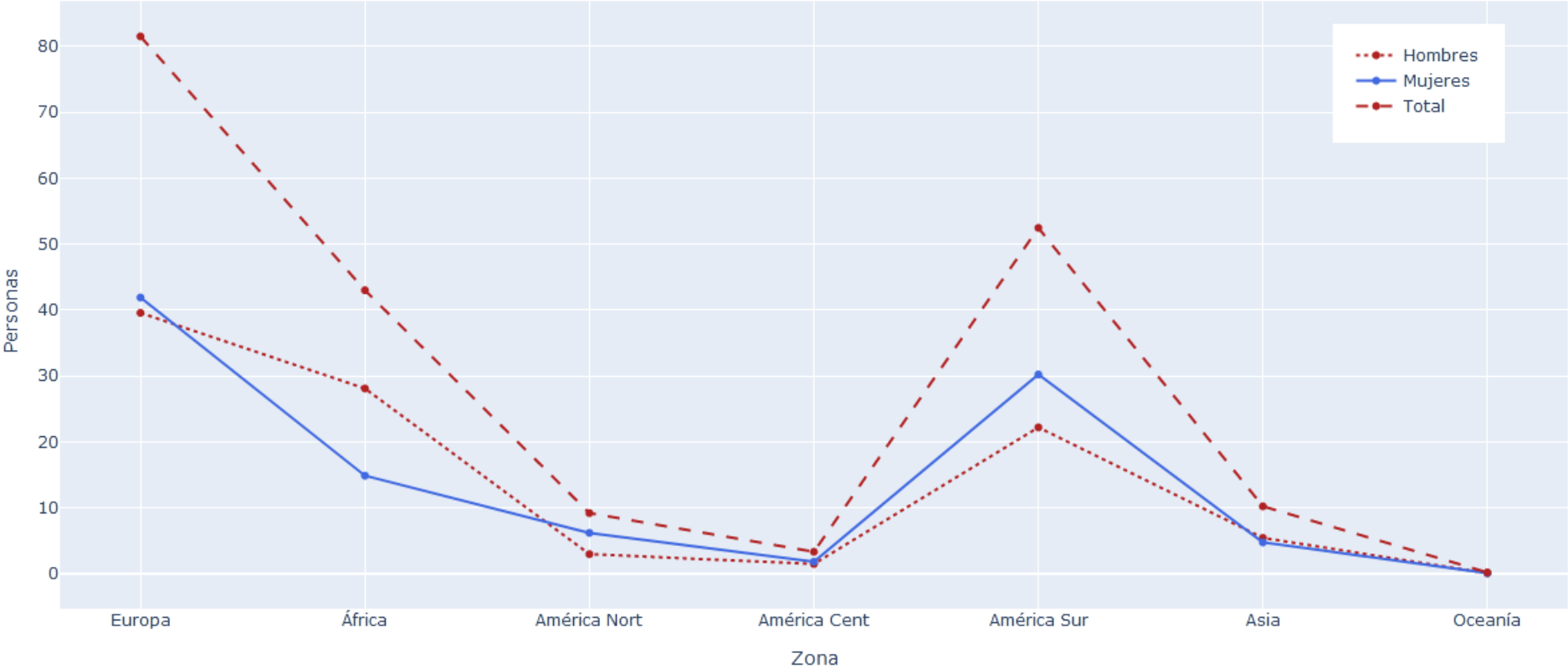
fig.show()

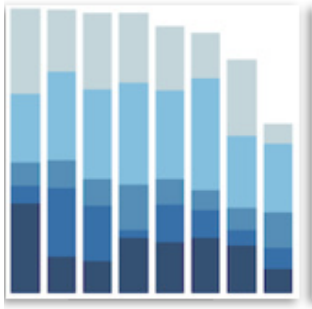
```

Comprueba que se realiza correctamente el gráfico ->
y luego cambia los colores y grosor de las 3 líneas.



Residentes por Zona origen datos 2001



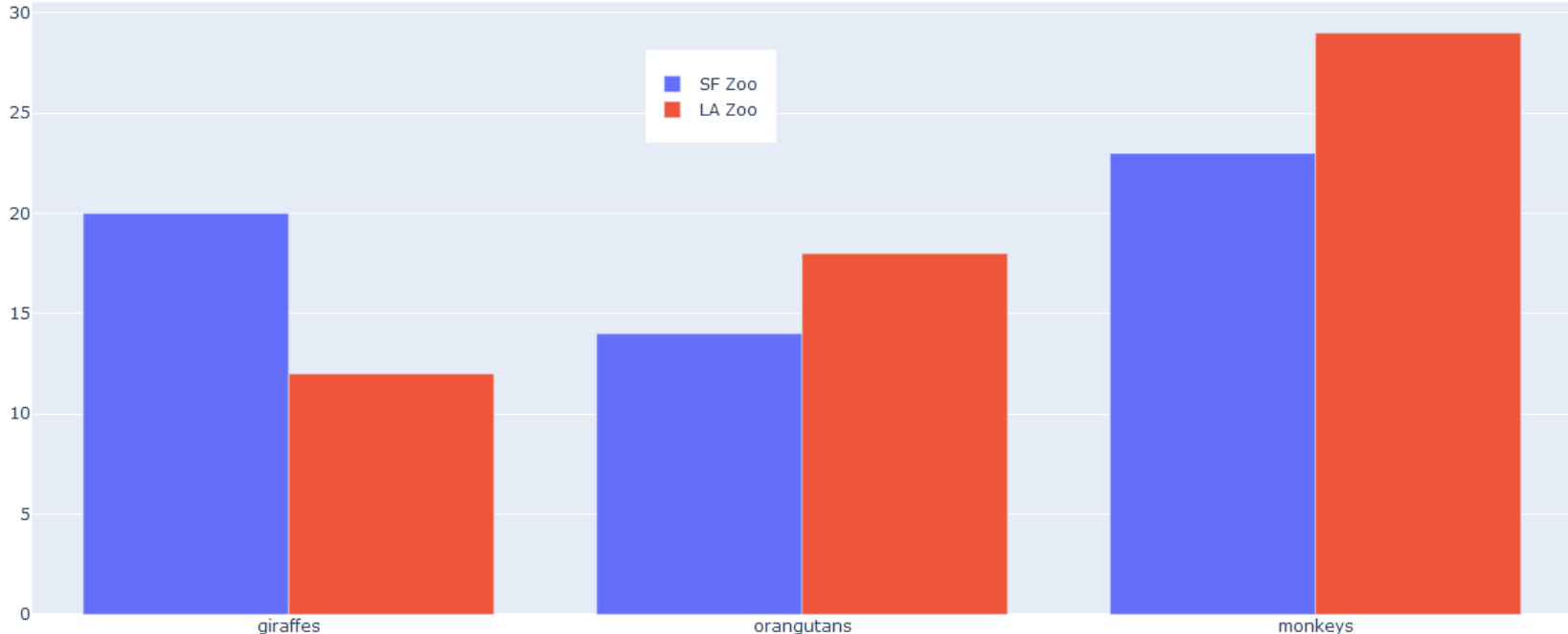


Bar Charts

2.2. Gráfico de Barras

```
import plotly.graph_objects as go
animals=['giraffes', 'orangutans', 'monkeys']

fig = go.Figure(data=[
    go.Bar(name='SF Zoo', x=animals, y=[20, 14, 23]),
    go.Bar(name='LA Zoo', x=animals, y=[12, 18, 29])]
# Change the bar mode
fig.update_layout(barmode='group')
fig.show()
```



Practica P02

Ejercicio 2. Pasa el gráfico de barras sobre animales a tu cuaderno-

Entra en la web de plotly <https://plotly.com/python/bar-charts/> y estudia como cambiar los colores.

Ejercicio 3. Puedes hacer un gráfico de barras con los datos de residentes por zona origen ?

Ejercicio 4. En la siguiente diapositiva hay un gráfico de tarta. Adáptalo para que muestre los datos de residentes por zona origen.

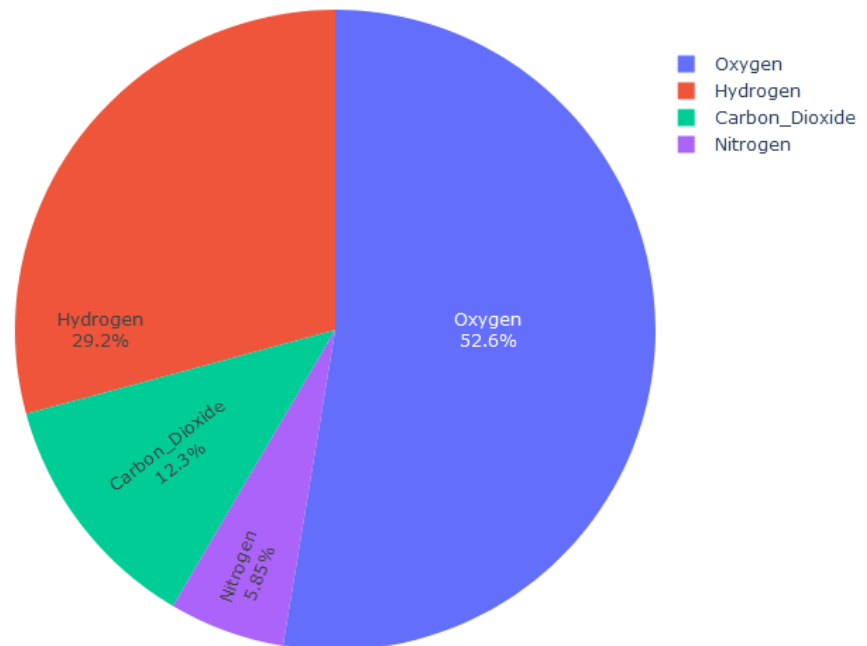
2.3. Tartas o gráficos circulares

```
import plotly.graph_objects as go
```

```
labels = ['Oxygen', 'Hydrogen', 'Carbon_Dioxide', 'Nitrogen']  
values = [4500, 2500, 1053, 500]
```

```
fig = go.Figure(data=[go.Pie(labels=labels, values=values,  
textinfo='label+percent', insidetextorientation='radial')])
```

```
fig.show()
```

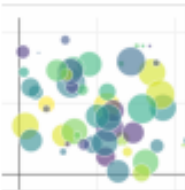


2.4. Otros gráficos básicos

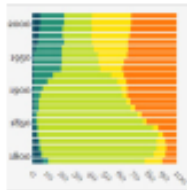
Plotly Fundamentals [🔗](#)



Dash -
Interactive
Python Apps



Static Image
Export



Updating
Plotly Graphs

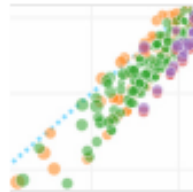


Jupyter
Notebook
Tutorial

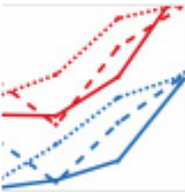


More Plotly
Fundamentals

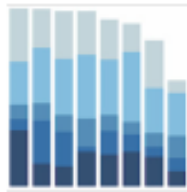
Basic Charts [🔗](#)



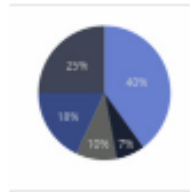
Scatter Plots



Line Charts



Bar Charts

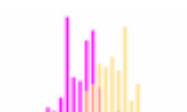
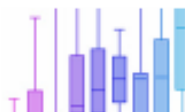
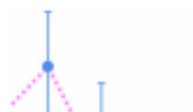


Pie Charts



More Basic
Charts

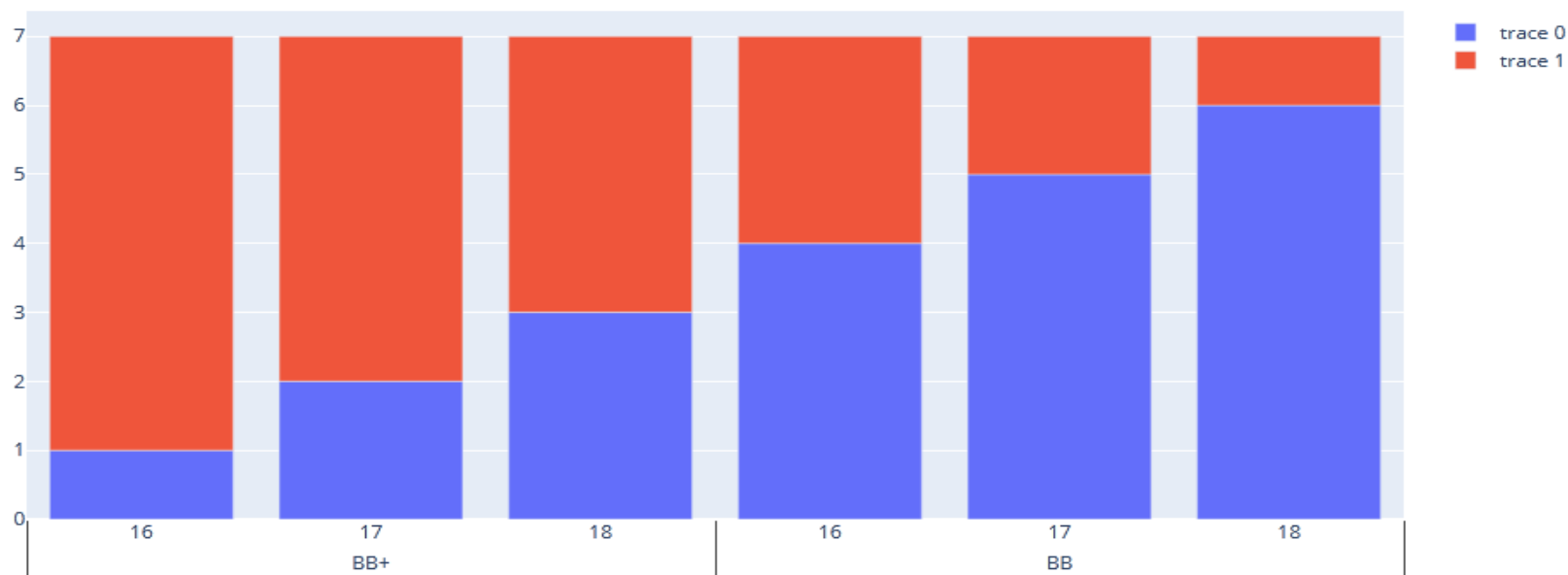
Statistical and Seaborn-style Charts [🔗](#)



Barras con ejes multicategoría

If your traces have arrays for `x` or `y`, then the axis type is automatically inferred to be `multicategory`

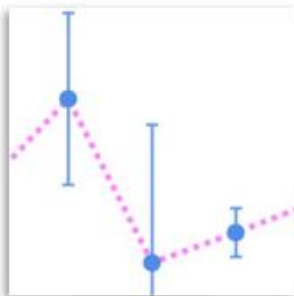
```
import plotly.graph_objects as go
x = [
    ["BB+", "BB+", "BB+", "BB", "BB", "BB"],
    [16, 17, 18, 16, 17, 18,]
]
fig = go.Figure()
fig.add_bar(x=x, y=[1, 2, 3, 4, 5, 6])
fig.add_bar(x=x, y=[6, 5, 4, 3, 2, 1])
fig.update_layout(barmode="relative")
fig.show()
```



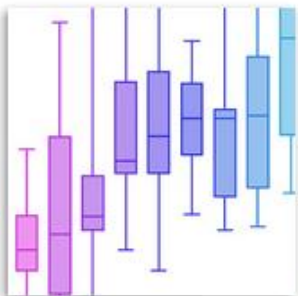
2.5. Gráficos estadísticos

Plotly Python Open Source Graphing Library Statistical Charts

Plotly's Python graphing library makes interactive, publication-quality graphs online. Examples of how to make statistical charts such as box plots, histograms, and distribution plots.



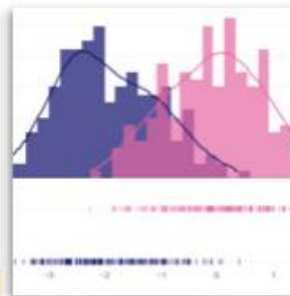
Error Bars



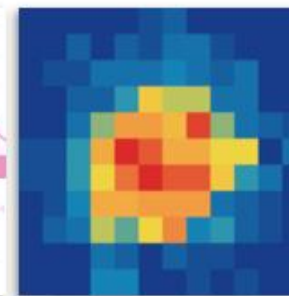
Box Plots



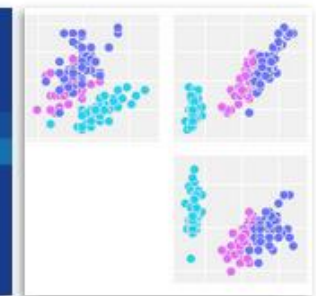
Histograms



Distplots



2D Histograms



Scatterplot Matrix

File Edit View Insert Cell Kernel Widgets Help



```
In [1]: import plotly.io as pio  
pio.renderers.default = "browser"
```

```
In [3]: import plotly.express as px  
data_canada = px.data.gapminder().query("country == 'Canada'")  
fig = px.bar(data_canada, x='year', y='pop')  
fig.show()
```

Si no ves bien los gráficos en jupyter:

```
import plotly.io as pio  
pio.renderers.default = "browser"
```

Practica P03

Ejercicio 1. Estudia y pasa a tu cuaderno 3 gráficos de la categoría Gráficos estadísticos de la web de Plotly.

Realiza cambios para comprobar que comprendes los parámetros principales.

<https://plotly.com/python/statistical-charts/>