

LENGUAJE PYTHON

Sesión No.4: Exploración y Análisis de datos

Jefferson A. Peña Torres

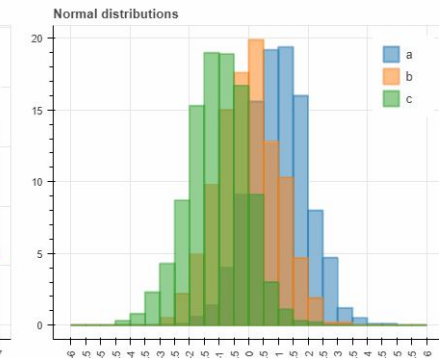
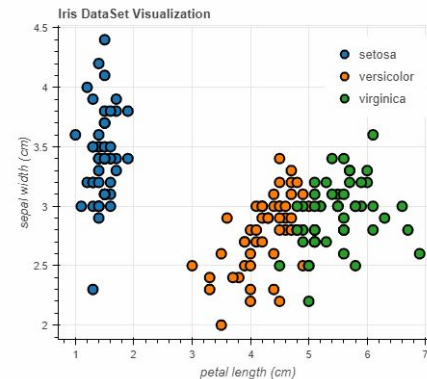
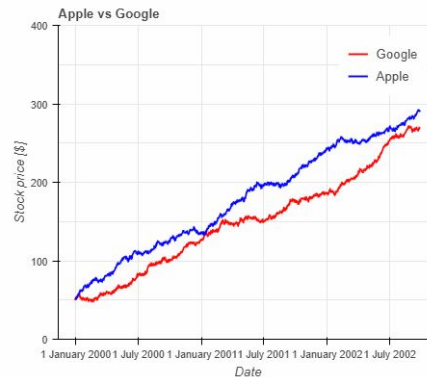
jefferson.amado.pena@correounivalle.edu.co

Escuela de Ingeniería de sistemas y computación (EISC)
Universidad del Valle, Cali
Enero 2020

Gráficos en Python

Las gráficas y visualizaciones son elementos fundamentales para entender, compartir y presentar información sobre los datos

- **Matplotlib:** Librería robusta para crear gráficas.
- **+ Seaborn:** Gráficos estadísticos
- **+ Folium:** Geolocalización basado en leaflet
- **+ Cartopy:** Cartografía
- **Bokeh:** Gráficas interactivas basadas en D3



Gráficos en Python

- Datos Numéricos

- replot
- scatterplot
- lineplot

- Datos Categórico

- catplot
- boxplot
- stripplot
- swarmplot
- etc

- Visualizar la distribución

- distplot
- kdeplot
- joinplot

- Regresión Lineal

- regplot
- lmlplot

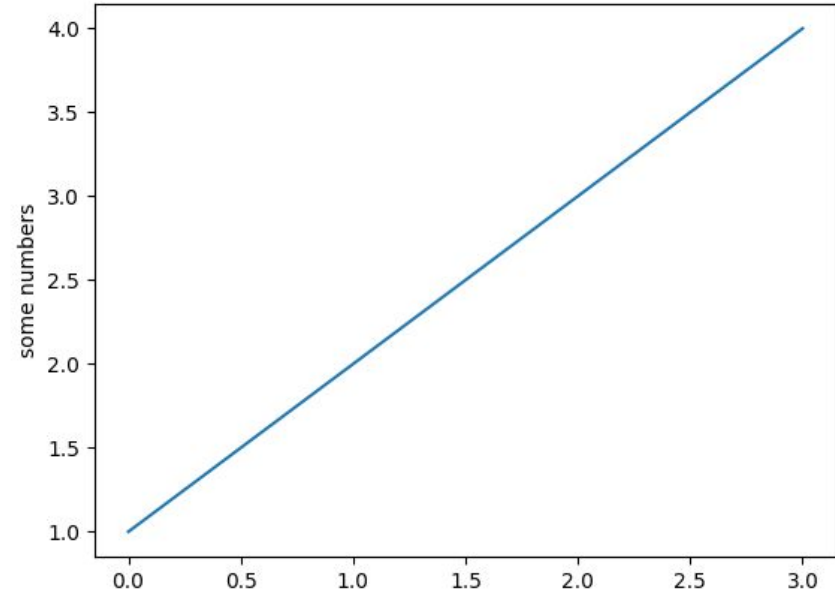
- Configuración

- **ejes**
- **titulos**
- **colores**

Gráficos en Python

Partes básicas de un gráfico con Matplotlib

```
import matplotlib.pyplot as plt
plt.plot([1, 2, 3, 4])
plt.ylabel('some numbers')
plt.show()
```

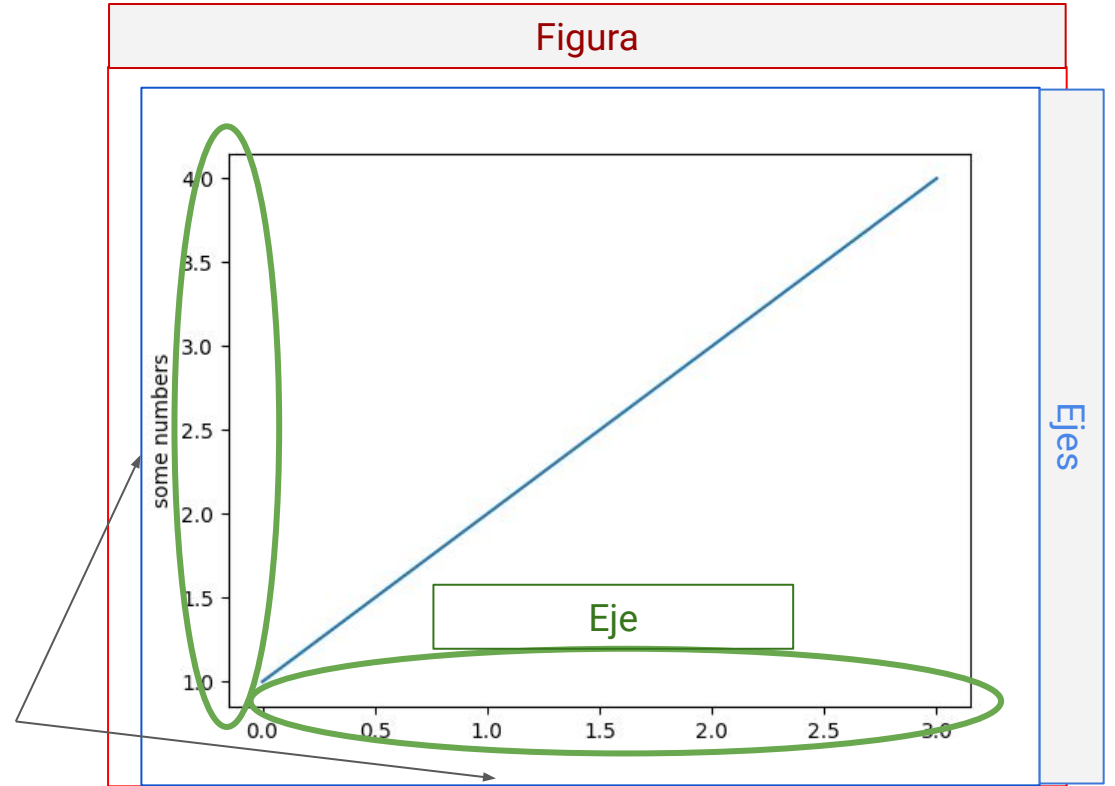


Gráficos en Python

Partes básicas de un gráfico con Matplotlib

```
import matplotlib.pyplot as plt  
plt.plot([1, 2, 3, 4])  
plt.ylabel('some numbers')  
plt.show()
```

Etiquetas



Gráficos en Python

Como ruta para graficar seguiremos estos pasos:

1. Crear figura
2. Obtener ejes
3. Dibujar sobre los ejes
4. Añadir elementos extras (anotaciones, títulos)
5. Render

```
import matplotlib.pyplot as plt
fig = plt.figure() # crear figura
ax = plt.subplot(1, 1, 1) # crear ejes
ax.plot([1,1]) # dibujar sobre los ejes
ax.plot([2,2]) # dibujar sobre los ejes
ax.plot([3,3]) # dibujar sobre los ejes
```

Gráficos en Python

Como ruta para graficar seguiremos estos pasos:

1. Crear figura
2. Obtener ejes
3. Dibujar sobre los ejes
4. Añadir elementos extras (anotaciones, títulos)
5. Render

```
N= 365
np.random.seed(2983)
time= pd.date_range('1/1/2020', periods=N)
y= 200 * (np.random.randn(N).cumsum()+40)

fig = plt.figure() # crear figura
ax = plt.subplot(1, 1, 1) # crear ejes
ax.plot(time, y) # dibujar sobre los ejes
```

Gráficos en Python

Ejercicio Básico

Un centro de reparación de vehículos ha realizado un conteo mensual por marcas para determinar cuántos vehículos por marca ha reparado en los últimos 5 meses

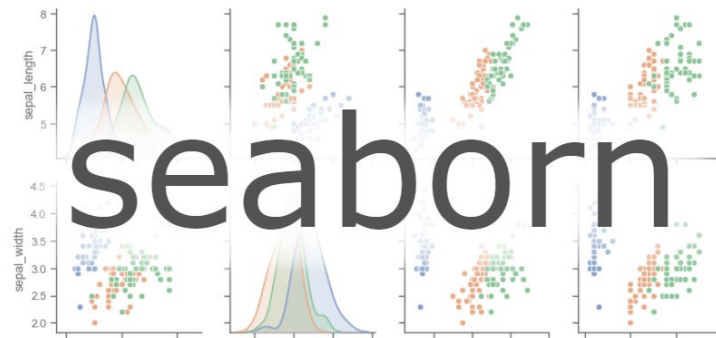
	CHEVROLET	BMW	M. BENZ	HONDA	HYUNDAI
SEP 19	50	34	45	30	30
OCT 19	40	23	30	43	18
NOV 19	70	12	20	56	20
DIC 19	32	10	16	98	56
ENERO 2020	89	18	10	45	10

Gráficos en Python

```
import seaborn as sns
import matplotlib.pyplot as plt
height = [62, 64, 69,
          75, 66, 68, 65,
          71, 76, 73]
weight = [120, 136, 148,
          175, 137, 165, 154,
          172, 200, 187]
```

```
sns.scatterplot(x=height, y=weight)
plt.show()
```

```
import seaborn as sns
import matplotlib.pyplot as plt
count_cars = ["Mazda" , "Mazda" , "Mazda" ,
              "Mazda" , "Dodge" , "Dodge" ,
              "Dodge" , "Dodge" , "Dodge" ,"Dodge"]
sns.countplot(x=count_cars)
plt.show()
```

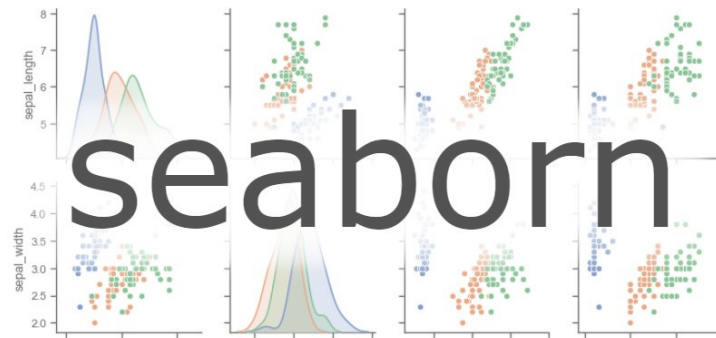


Gráficos en Python

```
import seaborn as sns
import matplotlib.pyplot as plt
```

```
x= np.random.normal(size=100)
sns.distplot(x)
plt.show()
```

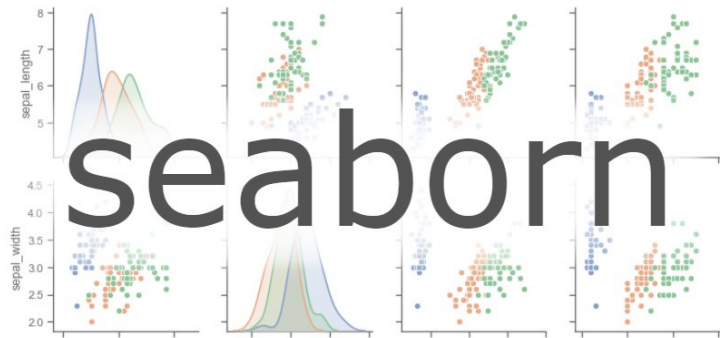
```
import seaborn as sns
import matplotlib.pyplot as plt
tips.head()
sns.regplot(x="total_bill", y="tip", data=tips)
plt.show()
```



Gráficos en Python

```
#%%  
# data for plot  
ts=pd.Series(np.random.randn(1000),  
             index=pd.date_range('1/1/2000',periods=1000))  
ts=ts.cumsum()  
df=pd.DataFrame(np.random.randn(1000,4),index=ts.index,  
               columns=['A','B','C','D'])  
  
df=df.cumsum()  
# bmh  
plt.style.use(['bmh'])  
plt.figure()  
df.plot()  
plt.legend(loc='best')
```

```
import seaborn as sns  
import matplotlib.pyplot as plt  
count_cars = ["Mazda" , "Mazda" , "Mazda" ,  
             "Mazda" , "Dodge" , "Dodge" ,  
             "Dodge" , "Dodge" , "Dodge" , "Dodge"]  
sns.countplot(x=count_cars)  
plt.show()
```



Gráficos en Python

```
from bokeh.io import output_file, show
from bokeh.plotting import figure
x = [1,2,3,4,5]
y = [8,6,5,2,3]
plot = figure()
plot.line(x, y, line_width=3)
plot.circle(x, y, fill_color='white', size=10)
output_file('basic.html')
show(plot)
```

```
##
# data for plot
from bokeh.io import output_file, show
from bokeh.plotting import figure
plot = figure(plot_width=400, tools='pan,box_zoom')
plot.circle([1,2,3,4,5], [8,6,5,2,3])
output_file('circle.html')
show(plot)
```



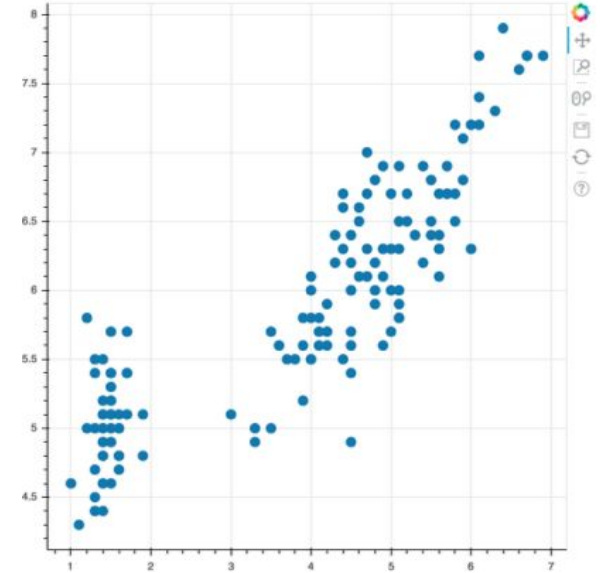
Bokeh

Gráficos en Python

```
from bokeh.io import output_file, show
from bokeh.plotting import figure
from bokeh.sampledata.iris import flowers
```

```
plot = figure()
plot.circle(flowers['petal_length'],
...: flowers['sepal_length'],
...: size=10)
```

```
output_file('pandas.html')
show(plot)
```



Bokeh



GRACIAS

Jefferson A. Peña Torres

jefferson.amado.pena@correounivalle.edu.co

Escuela de Ingeniería de sistemas y computación (EISC)
Universidad del Valle, Cali
Enero 2020