St Im # Im impo	1002S Herramientas computacionales: el arte de la analítica s a notebook with all your work for the final evidence of this course veles de dominio a demostrar con la evidencia G0202A preta interacciones entre variables relevantes en un problema como base para la construcción de modelos bivariados pasados en datos de un fenómeno investigado que le permita reproducir la respuesta del mismo. Es capaz de construir modelos bivariados que explique	ar .
# Imimpo	oreta interacciones entre variables relevantes en un problema, como base para la construcción de modelos bivariados basados en datos de un fenómeno investigado que le permita reproducir la respuesta del mismo. Es capaz de construir modelos bivariados que expliquendo contamiento de un fenómeno. Light information Name: Diego Sánchez Luna D: A00227831	en el
impo impo impo	My carreer: Ingeniería en Tecnologías Computacionales port libs port the packages that we will be using rt pandas as pd rt matplotlib.pyplot as plt rt seaborn as sns rt numpy as np	
# De RunI # If if R	sklearn.cluster import KMeans sklearn.datasets import load_digits fine where you are running the code: colab or local nColab = False # (False: no True: yes) running in colab: unInColab: # Mount your google drive in google colab from google.colab import drive drive.mount('/content/drive') # Find location	
else	#!Us "/content/drive/My Drive/Colab Notebooks/MachineLearningWithPython/" # Define path del proyecto Ruta = "/content/drive/My Drive/Colab Notebooks/MachineLearningWithPython/" : # Define path del proyecto Ruta = "C:/Users/GeekD/OneDrive/Documents/GitHub/TC1002S" URL = Ruta + "/Evidencia/a00227831.csv" Tt 1	
Us A1	ing your assigned dataset: Load data pd.read_csv(URL)	
Print	Data management the first 7 rows loc[0:7,:] named: 0	
2 3 4 5 6	2 -0.382497 1.130047 3 -0.319295 0.089919 4 -0.386405 1.624755 5 -0.749479 1.441339 6 -0.298645 0.152692	
df.t	ail() Unnamed: 0 x1 x2 995 0.058066 -0.180473 996 1.049360 0.128261 997 0.991189 0.298544 998 0.700671 -0.888454	
999 How prin	999 1.124124 -0.131446 many rows and columns are in your data? t("Columnas: "+str(df.shape[1])+"\t\tFilas: "+str(df.shape[0])) mas: 3 Filas: 1000 the name of all the columns	
Inde Wha	columns x(['Unnamed: 0', 'x1', 'x2'], dtype='object') tis the data type in each column types med: 0 int64	
dtyp Wha Print	e: object is the meaning of rows and columns? 1) ID del registro, se encuentra duplicada por la función de lectura de pandas 2) Medida flotante en la dimensión x1 3) Medida flotante en la dimensión x2 a statistical summary of your columns	
coun mean sto min	499.50000 0.252284 0.495796 288.819436 0.523188 0.889624 0.000000 -0.821863 -1.445373	
75% max Ansv	rer: 1) What is the minimum and maximum values of each variable? Max Min	
	unnamed: 99	
Rena	Se trata de los cuartiles de nuestro conjunto de datos y se distribuyen de la siguiente manera - 25% 25% de los datos es menor que o igual a este valor 50% La mediana. 50% de los datos es menor que o igual a este valor 75% 75% de los datos es menor que o igual a este valor. me the columns using the same name with capital letters df.rename(columns={"Unnamed: 0": "Index", "x1": "X1", "x2": "X2"})	
df =	me the columns to their original names df.rename(columns={"Index": "Unnamed: 0", "X1" : "x1", "X2" : "x2"}) wo different alternatives to get one of the columns t(df["x1"]) t('\n'+str(df.iloc[:,1])) 1.030877	
1 2 3 4 995 996 997 998 999 Name	1.050677 0.677208 -0.382497 -0.319295 -0.386405 0.058066 1.049360 0.991189 0.700671 1.124124 : x1, Length: 1000, dtype: float64	
01234995996997998999	1.030877 0.677208 -0.382497 -0.319295 -0.386405 0.058066 1.049360 0.991189 0.790671 1.124124	
Get a df.i	: x1, Length: 1000, dtype: float64 Islice of your dataset second and third columns and rows from 62 to 72 Ioc[62:72, 2:3] x2 0.076857 0.012296	
65 66 67 68 69	0.427920 2.148473 1.565853 2.153031 0.426566 0.309148 0.824856	
prin Valo	ne second and thrid columns, calculate the number of null and not null values and verify that their sum equals the total number of rows t("Valores nulos: "+str(df.iloc[:,2:3].isnull().sum())+ "\n\nValores no nulos: "+str(df.iloc[:,2:3].notnull().sum())) res nulos: x2	
dtyp Disca	res in nutus: X2 1000 e: int64 rrd the last column (Index in this case) rop("Unnamed: 0", axis=1, inplace = True) rtions d on the previous results, provide a description of your dataset.	
A3 Plot	response: te dataset se nos brindan dos columnas con datos de medición en el rango de los números reales en las dimensiones x1 y x2, dándonos los valores necesarios para graficar. Data visualization In the same figure the histogram of the two variables hist(df.x1, alpha = .3, 1s = "dotted", edgecolor = "black")	
plt. plt. plt.	X2	
Seperticiones 2		
Your Com Plot	d on this plots, provide a description of your data response here: p podemos observar en el histograma, los registros de la variable x1 tienen un rango mucho menor a aquellos encontrados en la variable x2. Además de que estas se concentran principalmente en los valores centrales entre -0.5 y 1.5. In the same figure the boxplot of the two variables	
plt. plt. plt. plt.	boxplot(data=df, orient="v") ylabel("Tamaño") xlabel("Variable") title("Mediciones") show() sers\GeekD\anaconda3\lib\site-packages\seaborn\categorical.py:82: FutureWarning: iteritems is deprecated and will be removed in a future version. Use .items instead. tot_data = [np.asarray(s, float) for k, s in iter_data] Mediciones Mediciones	
lamaño 		
plt. plt. plt.	Variable er plot of the two variables scatter(df.x1, df.x2, alpha = 0.6) title("Mediciones") xlabel("x1") ylabel("x2")	
Z (Mediciones .5 -	
A 2 Do K K =	response: Con estas gráficas volvemos a ver la diferencia que hay entre las dos mediciones, siendo x2 aquella con valores más altos. Dicho esto, podemos ver que forman una especie de figura con los valores en la gráfica de dispersión, dando a entender que existe un pat Kmeans means clusterin assuming a number of clusters according to your scatter plot. graphical de dispersión, dando a entender que existe un pat the properties of the	trón entre e
C:\U M_TH wa	ters = km.fit_predict(df) sers\GeekD\anaconda3\lib\site-packages\sklearn\cluster_kmeans.py:1382: UserWarning: KMeans is known to have a memory leak on Windows with MKL, when there are less chunks than available threads. You can avoid it by setting the environ READS=4. enings.warn(to your dataset a column with the assigned cluster to each data point Cluster"] = clusters	nment vari
df.C arra Print kmc kmc	the number associated to each cluster luster.unique() /([1, 0]) the centroids = km.cluster_centers_	
Print km.i	v([[-0.01879283, 1.21109641, 0.64970646],	
plt. plt. plt. plt. plt.	<pre>= df[df.Cluster==0] = df[df.Cluster==1] scatter(df1.x1,df1.x2, label = "Cluster 0", alpha = 0.5) scatter(df2.x1, df2.x2, label = "Cluster 1", alpha = 0.5) scatter(kmc[:,0],kmc[:,1], marker = "X", color = "Purple", s = 200) title("Mediciones con cluster asignado y centroides") xlabel("x1") ylabel("x2") legend()</pre>	
plt.	Mediciones con cluster asignado y centroides Cluster 0 Cluster 1	
-0 -1 -1 Que :	100.5	
En e	de a detailed description of your results ta caso, dividiendolo en dos grupos podemos ver como la figura es dividida a la mitad de manera horizontal en vez de separar ambas secciones de la figura. Esto se debe en primer lugar a la aleatoriedad que presenta la función de KMeans así como la posición de los cent Elbow Plot pute the elbow plot ### List of sse	ntroides asig
Com	<pre>= [] ine K [1,2,3,4,5,6,7,8,9,10]</pre>	
#Ini SSE #Def K = #For for	SSE.append(tempkm.inertia_) sers\GeekD\anaconda3\lib\site-packages\sklearn\cluster_kmeans.py:1382: UserWarning: KMeans is known to have a memory leak on Windows with MKL, when there are less chunks than available threads. You can avoid it by setting the environ READS=4.	nment vari
#Ini SSE #Def K = #For for C:\U M_TH wa C:\U M_TH wa C:\U M_TH wa C:\U M_TH wa C:\U M_TH	SEE.append(tempkm.inertia_) Sers\GeekD\anaconda3\lib\site-packages\sklearn\cluster_kmeans.py:1382: UserWarning: KMeans is known to have a memory leak on Windows with MKL, when there are less chunks than available threads. You can avoid it by setting the environ READS=4. **Prings.warn(SEADS=4. **Pr	nment vari nment vari nment vari
#Ini SSE #Def K = #For for C:\U M_TH wa C:\U	SSE.append(tempkm.inertia_) Sers\GeekD\anaconda3\lib\site-packages\sklearn\cluster_kmeans.py:1382: UserWarning: KMeans is known to have a memory leak on Windows with MKL, when there are less chunks than available threads. You can avoid it by setting the environ raings.warn(Sers\GeekD\anaconda3\lib\site-packages\sklearn\cluster_kmeans.py:1382: UserWarning: KMeans is known to have a memory leak on Windows with MKL, when there are less chunks than available threads. You can avoid it by setting the environ raings.warn(Sers\GeekD\anaconda3\lib\site-packages\sklearn\cluster_kmeans.py:1382: UserWarning: KMeans is known to have a memory leak on Windows with MKL, when there are less chunks than available threads. You can avoid it by setting the environ raings.warn(Sers\GeekD\anaconda3\lib\site-packages\sklearn\cluster_kmeans.py:1382: UserWarning: KMeans is known to have a memory leak on Windows with MKL, when there are less chunks than available threads. You can avoid it by setting the environ raings.warn(Sers\GeekD\anaconda3\lib\site-packages\sklearn\cluster_kmeans.py:1382: UserWarning: KMeans is known to have a memory leak on Windows with MKL, when there are less chunks than available threads. You can avoid it by setting the environ raings.warn(Sers\GeekD\anaconda3\lib\site-packages\sklearn\cluster_kmeans.py:1382: UserWarning: KMeans is known to have a memory leak on Windows with MKL, when there are less chunks than available threads. You can avoid it by setting the environ raings.warn(Sers\GeekD\anaconda3\lib\site-packages\sklearn\cluster_kmeans.py:1382: UserWarning: KMeans is known to have a memory leak on Windows with MKL, when there are less chunks than available threads. You can avoid it by setting the environ raings. We have the packages and the packages are less chunks than available threads. You can avoid it by setting the environ raings. We have the packages are less chunks than available threads. You can avoid it by setting the environ raings. We have the packages are less chunks than ava	nment vari nment vari nment vari nment vari nment vari
#Ini SSE #Def K = #For for C:\UM_TH	Ser-Speedo\anaconda3\lib\site-packages\sklearn\cluster\kmeans.py:1382: UserWarning: KMeans is known to have a memory leak on Windows with MKL, when there are less chunks than available threads. You can avoid it by setting the environ RADS-4. "nings.warn(ser-Speeko\anaconda3\lib\site-packages\sklearn\cluster\kmeans.py:1382: UserWarning: KMeans is known to have a memory leak on Windows with MKL, when there are less chunks than available threads. You can avoid it by setting the environ RADS-4. "nings.warn(ser-Speeko\anaconda3\lib\site-packages\sklearn\cluster\kmeans.py:1382: UserWarning: KMeans is known to have a memory leak on Windows with MKL, when there are less chunks than available threads. You can avoid it by setting the environ RADS-4. "nings.warn(ser-Speeko\anaconda3\lib\site-packages\sklearn\cluster\kmeans.py:1382: UserWarning: KMeans is known to have a memory leak on Windows with MKL, when there are less chunks than available threads. You can avoid it by setting the environ RADS-4. "nings.warn(ser-Speeko\anaconda3\lib\site-packages\sklearn\cluster\kmeans.py:1382: UserWarning: KMeans is known to have a memory leak on Windows with MKL, when there are less chunks than available threads. You can avoid it by setting the environ RADS-4. "nings.warn(ser-Speeko\anaconda3\lib\site-packages\sklearn\cluster\kmeans.py:1382: UserWarning: KMeans is known to have a memory leak on Windows with MKL, when there are less chunks than available threads. You can avoid it by setting the environ RADS-4. "nings.warn(ser-Speeko\anaconda3\lib\site-packages\sklearn\cluster\kmeans.py:1382: UserWarning: KMeans is known to have a memory leak on Windows with MKL, when there are less chunks than available threads. You can avoid it by setting the environ RADS-4. "nings.warn(ser-Speeko\anaconda3\lib\site-packages\sklearn\cluster\kmeans.py:1382: UserWarning: KMeans is known to have a memory leak on Windows with MKL, when there are less chunks than available threads. You can avoid it by setting the environ RADS-4. "nings.warn	nment vari nment vari nment vari nment vari nment vari nment vari
#Ini SSE #Def K = #For C:\UM_TH	Sees_SeedDynacondu3\lib\site-packages\sklearm\cluster_kmeans.py:1382: UserNaming: KMeans is known to have a memory leak on Windows with NKL, when there are less chunks than available threads. You can avoid it by setting the environ tacks. **Tings.warn(**ERASS-4.** **Tings.warn(**ERASS-4.** **Tings.warn(**Tings	nment vari nment vari nment vari nment vari nment vari nment vari
#Ini SSE #Def K = #For for C:\UM_TH wa C:\UM_TM wa C:	For executive training to the control of the contro	nment vari nment vari nment vari nment vari nment vari nment vari
#Ini SSE #Def K = #For for C:\UM_TH wa C:\UM_TM wa C:	Site appel (Control Control Co	nment vari nment vari nment vari nment vari nment vari nment vari
# Init # SSE # For C:\U M	The product of part of the product o	nment vari nment vari nment vari nment vari nment vari nment vari
## PL plt. 1200 1000 1000 1000 1000 1000 1000 10	Comparison Com	nment vari nment vari nment vari nment vari nment vari nment vari
### C:\UM_SE ### For for M_SE ### For for for M_SE ### For for for M_SE ### For for		nment vari nment vari nment vari nment vari nment vari nment vari
## C M	The control of the co	nment vari nment vari nment vari nment vari nment vari nment vari
## Company	## Company of Company	nment vari nment vari nment vari nment vari nment vari nment vari
Com #Ini Sign #For C:\Um Sign #For C:\	Company Comp	nment varianment varianment varianment varianment varianment varianment varia
Com #Ini SEP #For for C:\U M_SEP #For for C:\U M_SEP #For C:\U M_SEP	Company Comp	nment vari nment vari nment vari nment vari nment vari nment vari
Com #5 #5 #5 #6 #6 #6 #6 #6 #6 #6 #6 #6 #6 #6 #6 #6	A CONTROLLEY CONTROLLE	nment vari nment vari nment vari nment vari nment vari nment vari
Community and a second control of the second	## Company of the Com	nment vari nment vari nment vari nment vari nment vari
Compared and a second and a sec	Company of the comp	nment vari
Community of the service of the serv		nment vari
Common the second parameter of	The content of the co	nment varianment varia
Community of the commun		nment varianment varia