

	VICERRECTORADO DOCENTE	Código: GUIA-PRL-001
	CONSEJO ACADÉMICO	Aprobación: 2016/04/06
Formato: Guía de Práctica de Laboratorio / Talleres / Centros de Simulación		

		FORMATO DE INFORME DE PRÁCTICA DE LABORATORIO / TALLERES / CENTROS DE SIMULACIÓN – PARA ESTUDIANTES																																																													
CARRERA: Computación		ASIGNATURA: Simulación																																																													
NRO. PRÁCTICA:		TÍTULO PRÁCTICA: Regresión COVID																																																													
OBJETIVO ALCANZADO: Comprensión sobre la práctica de la regresión que se puede realizar a un conjunto de datos para obtener predicciones.																																																															
ACTIVIDADES DESARROLLADAS																																																															
1. Empezamos por importar las librerías necesarias para el Desarrollo de la práctica																																																															
<pre>import pandas as pd import numpy as np from matplotlib import pyplot as plt from datetime import datetime, timedelta from sklearn.model_selection import train_test_split as tts from sklearn.linear_model import LinearRegression #from sklearn.pipeline import Pipeline from sklearn.metrics import mean_absolute_error</pre>																																																															
2. Procedemos con la carga y procesamiento de los datos																																																															
<pre>data = pd.read_csv('owid-covid-data.csv') dataF = data dataF = dataF.loc[:, ['location', 'date', 'total_cases', 'new_cases']] dataF = dataF[dataF['location']=='Ecuador'] print(dataF)</pre> <table><thead><tr><th></th><th>location</th><th>date</th><th>total_cases</th><th>new_cases</th></tr></thead><tbody><tr><td>22684</td><td>Ecuador</td><td>2020-03-01</td><td>6.0</td><td>6.0</td></tr><tr><td>22685</td><td>Ecuador</td><td>2020-03-02</td><td>6.0</td><td>0.0</td></tr><tr><td>22686</td><td>Ecuador</td><td>2020-03-03</td><td>7.0</td><td>1.0</td></tr><tr><td>22687</td><td>Ecuador</td><td>2020-03-04</td><td>10.0</td><td>3.0</td></tr><tr><td>22688</td><td>Ecuador</td><td>2020-03-05</td><td>13.0</td><td>3.0</td></tr><tr><td>...</td><td>...</td><td>...</td><td>...</td><td>...</td></tr><tr><td>23109</td><td>Ecuador</td><td>2021-04-30</td><td>381862.0</td><td>1173.0</td></tr><tr><td>23110</td><td>Ecuador</td><td>2021-05-01</td><td>384589.0</td><td>2727.0</td></tr><tr><td>23111</td><td>Ecuador</td><td>2021-05-02</td><td>387299.0</td><td>2710.0</td></tr><tr><td>23112</td><td>Ecuador</td><td>2021-05-03</td><td>388046.0</td><td>747.0</td></tr><tr><td>23113</td><td>Ecuador</td><td>2021-05-04</td><td>391242.0</td><td>3196.0</td></tr></tbody></table> <p>[430 rows x 4 columns]</p>					location	date	total_cases	new_cases	22684	Ecuador	2020-03-01	6.0	6.0	22685	Ecuador	2020-03-02	6.0	0.0	22686	Ecuador	2020-03-03	7.0	1.0	22687	Ecuador	2020-03-04	10.0	3.0	22688	Ecuador	2020-03-05	13.0	3.0	23109	Ecuador	2021-04-30	381862.0	1173.0	23110	Ecuador	2021-05-01	384589.0	2727.0	23111	Ecuador	2021-05-02	387299.0	2710.0	23112	Ecuador	2021-05-03	388046.0	747.0	23113	Ecuador	2021-05-04	391242.0	3196.0
	location	date	total_cases	new_cases																																																											
22684	Ecuador	2020-03-01	6.0	6.0																																																											
22685	Ecuador	2020-03-02	6.0	0.0																																																											
22686	Ecuador	2020-03-03	7.0	1.0																																																											
22687	Ecuador	2020-03-04	10.0	3.0																																																											
22688	Ecuador	2020-03-05	13.0	3.0																																																											
...																																																											
23109	Ecuador	2021-04-30	381862.0	1173.0																																																											
23110	Ecuador	2021-05-01	384589.0	2727.0																																																											
23111	Ecuador	2021-05-02	387299.0	2710.0																																																											
23112	Ecuador	2021-05-03	388046.0	747.0																																																											
23113	Ecuador	2021-05-04	391242.0	3196.0																																																											
3. Graficamos los datos obtenidos																																																															
																																																															

4. Le damos el formato adecuado a las fechas para poder realizar correctamente la predicción con la regresión

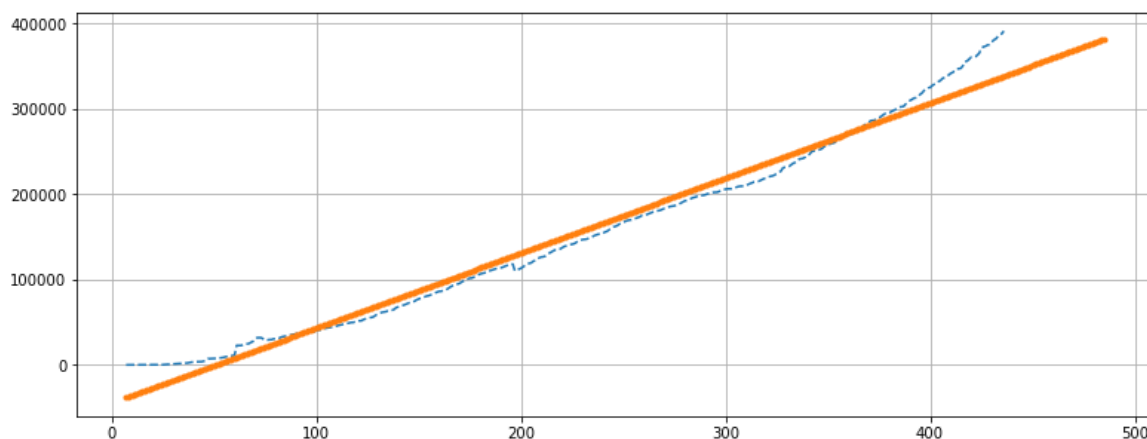
```
dataF['date'] = dataF['date'].map(lambda x: (datetime.strptime(x, '%Y-%m-%d') - datetime.strptime("2020-02-23", '%Y-%m-%d')).days)
```

5. Creamos las variables necesarias para realizar el entrenamiento

```
xtrain, xtest, ytrain, ytest = tts(dataF['date'], dataF['total_cases'], test_size=0.2)
modelo = LinearRegression()
entrenamiento = np.array(xtrain).reshape(-1, 1)
modelo.fit(entrenamiento, ytrain)
```

6. Por último, ejecutamos la regresión lineal y graficamos

```
prediccion = list(range(min(dataF['date']), max(dataF['date'])+50))
pred = modelo.predict(np.array(prediccion).reshape(-1, 1))
plt.figure(figsize=(13,5))
plt.grid(True)
plt.plot(dataF['date'], dataF['total_cases'], '--', label='Datos Reales')
plt.plot(prediccion, pred, '.', label="Regresión Lineal")
plt.show()
```



RESULTADO(S) OBTENIDO(S): Una predicción por regresión lineal de los datos existentes con respecto al covid

CONCLUSIONES: podemos observar de manera práctica y clara como se realiza una regresión lineal que nos permite observar una predicción, en este caso, bastante decente.

Nombre de los estudiantes: Alejandro Enríquez