

# Universidad Politécnica Salesiana

Nombre: Jéssica Ñauta

Materia: Simulación

## Covid-19 infección en Ecuador. Modelos matemáticos y predicciones

Una comparación de modelos, lineal, polinómico, logísticos y exponenciales aplicados a la infección por el virus Covid-19

Se realiza un análisis matemático simple del crecimiento de la infección en Python y dos modelos para comprender mejor la evolución de la infección.

Se crean modelos de series temporales del número total de personas infectadas hasta la fecha (es decir, las personas realmente infectadas más las personas que han sido infectadas). Estos modelos tienen parámetros, que se estimarán por ajuste de curva.

```
In [3]: # Importar las librerías para el análisis
import pandas as pd
import numpy as np
from datetime import datetime, timedelta
from sklearn.metrics import mean_squared_error
from scipy.optimize import curve_fit
from scipy.optimize import fsolve
from sklearn import linear_model
import matplotlib.pyplot as plt
%matplotlib inline
```

```
In [4]: # Actualizar los datos (URL)
# Datos obtenidos de: https://github.com/owid/covid-19-data/tree/master/public/data
url = 'owid-covid-data.csv'
df = pd.read_csv(url)
df = df.fillna(0)
df
```

```
Out[4]:
```

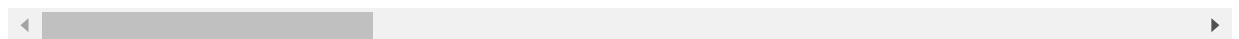
	iso_code	continent	location	date	total_cases	new_cases	new_cases_smoothed	total_
0	AFG	Asia	Afghanistan	2019-12-31	0.0	0.0	0.0	
1	AFG	Asia	Afghanistan	2020-01-01	0.0	0.0	0.0	
2	AFG	Asia	Afghanistan	2020-01-02	0.0	0.0	0.0	
3	AFG	Asia	Afghanistan	2020-01-03	0.0	0.0	0.0	
4	AFG	Asia	Afghanistan	2020-01-04	0.0	0.0	0.0	
5	AFG	Asia	Afghanistan	2020-01-05	0.0	0.0	0.0	

	iso_code	continent	location	date	total_cases	new_cases	new_cases_smoothed	total_
6	AFG	Asia	Afghanistan	2020-01-06	0.0	0.0	0.0	
7	AFG	Asia	Afghanistan	2020-01-07	0.0	0.0	0.0	
8	AFG	Asia	Afghanistan	2020-01-08	0.0	0.0	0.0	
9	AFG	Asia	Afghanistan	2020-01-09	0.0	0.0	0.0	
10	AFG	Asia	Afghanistan	2020-01-10	0.0	0.0	0.0	
11	AFG	Asia	Afghanistan	2020-01-11	0.0	0.0	0.0	
12	AFG	Asia	Afghanistan	2020-01-12	0.0	0.0	0.0	
13	AFG	Asia	Afghanistan	2020-01-13	0.0	0.0	0.0	
14	AFG	Asia	Afghanistan	2020-01-14	0.0	0.0	0.0	
15	AFG	Asia	Afghanistan	2020-01-15	0.0	0.0	0.0	
16	AFG	Asia	Afghanistan	2020-01-16	0.0	0.0	0.0	
17	AFG	Asia	Afghanistan	2020-01-17	0.0	0.0	0.0	
18	AFG	Asia	Afghanistan	2020-01-18	0.0	0.0	0.0	
19	AFG	Asia	Afghanistan	2020-01-19	0.0	0.0	0.0	
20	AFG	Asia	Afghanistan	2020-01-20	0.0	0.0	0.0	
21	AFG	Asia	Afghanistan	2020-01-21	0.0	0.0	0.0	
22	AFG	Asia	Afghanistan	2020-01-22	0.0	0.0	0.0	
23	AFG	Asia	Afghanistan	2020-01-23	0.0	0.0	0.0	
24	AFG	Asia	Afghanistan	2020-01-24	0.0	0.0	0.0	
25	AFG	Asia	Afghanistan	2020-01-25	0.0	0.0	0.0	
26	AFG	Asia	Afghanistan	2020-01-26	0.0	0.0	0.0	
27	AFG	Asia	Afghanistan	2020-01-27	0.0	0.0	0.0	
28	AFG	Asia	Afghanistan	2020-01-28	0.0	0.0	0.0	

	iso_code	continent	location	date	total_cases	new_cases	new_cases_smoothed	total_
29	AFG	Asia	Afghanistan	2020-01-29	0.0	0.0	0.0	
...	...	...	...	...	...	...	...	
56081	0	0	International	2020-10-13	696.0	0.0	0.0	
56082	0	0	International	2020-10-14	696.0	0.0	0.0	
56083	0	0	International	2020-10-15	696.0	0.0	0.0	
56084	0	0	International	2020-10-16	696.0	0.0	0.0	
56085	0	0	International	2020-10-17	696.0	0.0	0.0	
56086	0	0	International	2020-10-18	696.0	0.0	0.0	
56087	0	0	International	2020-10-19	696.0	0.0	0.0	
56088	0	0	International	2020-10-20	696.0	0.0	0.0	
56089	0	0	International	2020-10-21	696.0	0.0	0.0	
56090	0	0	International	2020-10-22	696.0	0.0	0.0	
56091	0	0	International	2020-10-23	696.0	0.0	0.0	
56092	0	0	International	2020-10-24	696.0	0.0	0.0	
56093	0	0	International	2020-10-25	696.0	0.0	0.0	
56094	0	0	International	2020-10-26	696.0	0.0	0.0	
56095	0	0	International	2020-10-27	696.0	0.0	0.0	
56096	0	0	International	2020-10-28	696.0	0.0	0.0	
56097	0	0	International	2020-10-29	696.0	0.0	0.0	
56098	0	0	International	2020-10-30	696.0	0.0	0.0	
56099	0	0	International	2020-10-31	696.0	0.0	0.0	
56100	0	0	International	2020-11-01	696.0	0.0	0.0	
56101	0	0	International	2020-11-02	696.0	0.0	0.0	

	iso_code	continent	location	date	total_cases	new_cases	new_cases_smoothed	total_
<b>56102</b>	0	0	International	2020-11-03	696.0	0.0	0.0	
<b>56103</b>	0	0	International	2020-11-04	696.0	0.0	0.0	
<b>56104</b>	0	0	International	2020-11-05	696.0	0.0	0.0	
<b>56105</b>	0	0	International	2020-11-06	696.0	0.0	0.0	
<b>56106</b>	0	0	International	2020-11-07	696.0	0.0	0.0	
<b>56107</b>	0	0	International	2020-11-08	696.0	0.0	0.0	
<b>56108</b>	0	0	International	2020-11-09	696.0	0.0	0.0	
<b>56109</b>	0	0	International	2020-11-10	696.0	0.0	0.0	
<b>56110</b>	0	0	International	2020-11-11	696.0	0.0	0.0	

56111 rows × 49 columns



Imprimos los resultados y agregamos el numero del dia

```
In [5]: df = df[df['location'].isin(['Ecuador'])] #Filtro la Informacion solo para Ecuador
df = df.loc[:,['date','total_cases']] #Selecciono las columnas de analisis
# Expresar las fechas en numero de dias desde el 01 Enero
FMT = '%Y-%m-%d'
date = df['date']
df['date'] = date.map(lambda x : (datetime.strptime(x, FMT) - datetime.strptime("2020-01-01", FMT)).days)
df
```

Out[5]:

	date	total_cases
<b>14965</b>	-1	0.0
<b>14966</b>	0	0.0
<b>14967</b>	1	0.0
<b>14968</b>	2	0.0
<b>14969</b>	3	0.0
<b>14970</b>	4	0.0
<b>14971</b>	5	0.0
<b>14972</b>	6	0.0
<b>14973</b>	7	0.0
<b>14974</b>	8	0.0
<b>14975</b>	9	0.0
<b>14976</b>	10	0.0

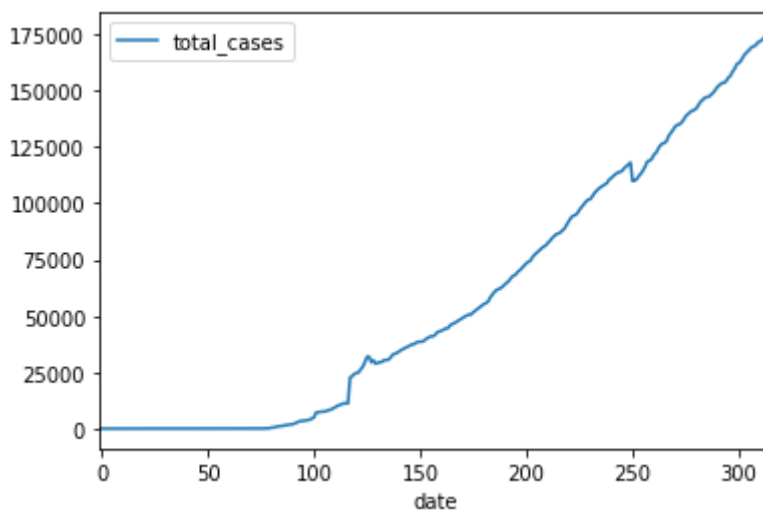
	date	total_cases
<b>14977</b>	11	0.0
<b>14978</b>	12	0.0
<b>14979</b>	13	0.0
<b>14980</b>	14	0.0
<b>14981</b>	15	0.0
<b>14982</b>	16	0.0
<b>14983</b>	17	0.0
<b>14984</b>	18	0.0
<b>14985</b>	19	0.0
<b>14986</b>	20	0.0
<b>14987</b>	21	0.0
<b>14988</b>	22	0.0
<b>14989</b>	23	0.0
<b>14990</b>	24	0.0
<b>14991</b>	25	0.0
<b>14992</b>	26	0.0
<b>14993</b>	27	0.0
<b>14994</b>	28	0.0
...	...	...
<b>15252</b>	286	147315.0
<b>15253</b>	287	148171.0
<b>15254</b>	288	149083.0
<b>15255</b>	289	150360.0
<b>15256</b>	290	151659.0
<b>15257</b>	291	152422.0
<b>15258</b>	292	153289.0
<b>15259</b>	293	153423.0
<b>15260</b>	294	154115.0
<b>15261</b>	295	155625.0
<b>15262</b>	296	156451.0
<b>15263</b>	297	158270.0
<b>15264</b>	298	159614.0
<b>15265</b>	299	161635.0
<b>15266</b>	300	162178.0
<b>15267</b>	301	163192.0
<b>15268</b>	302	164908.0

	date	total_cases
15269	303	166302.0
15270	304	167147.0
15271	305	168192.0
15272	306	169194.0
15273	307	169562.0
15274	308	170110.0
15275	309	171433.0
15276	310	171783.0
15277	311	172508.0
15278	312	173486.0
15279	313	174907.0
15280	314	175269.0
15281	315	175711.0

317 rows × 2 columns

```
In [6]: df.plot(x='date', y='total_cases')
```

```
Out[6]: <AxesSubplot:xlabel='date'>
```



Ahora podemos analizar los cuatro modelos que tomaré en el examen, que son la función lineal, polinómica, logística y la función exponencial. Cada modelo tiene tres parámetros, que se estimarán mediante un cálculo de ajuste de curva en los datos históricos.

## EL modelo lineal

La regresión lineal es un algoritmo de aprendizaje supervisado que se utiliza en Machine Learning y en estadística. En su versión más sencilla, lo que haremos es «dibujar una recta» que nos indicará la tendencia de un conjunto de datos continuos.

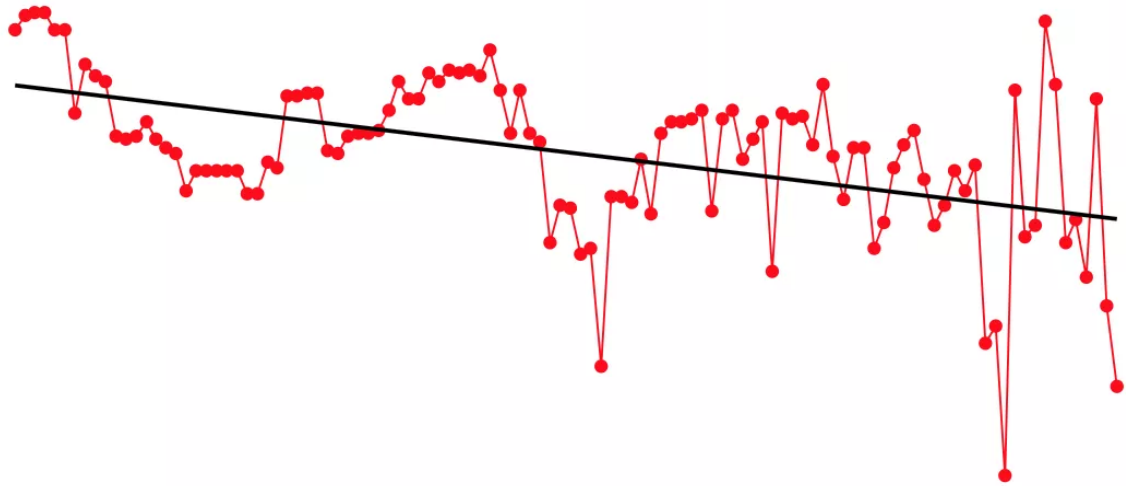
Recordemos rápidamente la fórmula de la recta:

$$Y = mX + b$$

Donde Y es el resultado, X es la variable, m la pendiente (o coeficiente) de la recta y b la constante o también conocida como el «punto de corte con el eje Y» en la gráfica (cuando X=0)

Ejemplo

## The development in Pizza prices in Denmark from 2009 to 2018



Recordemos que los algoritmos de Machine Learning Supervisados, aprenden por sí mismos y - en este caso- a obtener automáticamente esa «recta» que buscamos con la tendencia de predicción. Para hacerlo se mide el error con respecto a los puntos de entrada y el valor «Y» de salida real.

```
In [7]: x = list(df.iloc[:, 0]) # Fecha
y = list(df.iloc[:, 1]) # Numero de casos
# Creamos el objeto de Regresión Lineal
regr = linear_model.LinearRegression()

# Entrenamos nuestro modelo
regr.fit(np.array(x).reshape(-1, 1), y)

# Veamos los coeficientes obtenidos, En nuestro caso, serán la tangente
print('Coefficients: \n', regr.coef_)
# Este es el valor donde corta el eje Y (en X=0)
print('Independent term: \n', regr.intercept_)
# Error Cuadrado Medio
```

```
Coefficients:
[593.2518536]
Independent term:
-35652.30126823374
```

De la ecuación de la recta  $y = mX + b$  nuestra pendiente «m» es el coeficiente y el término independiente «b»

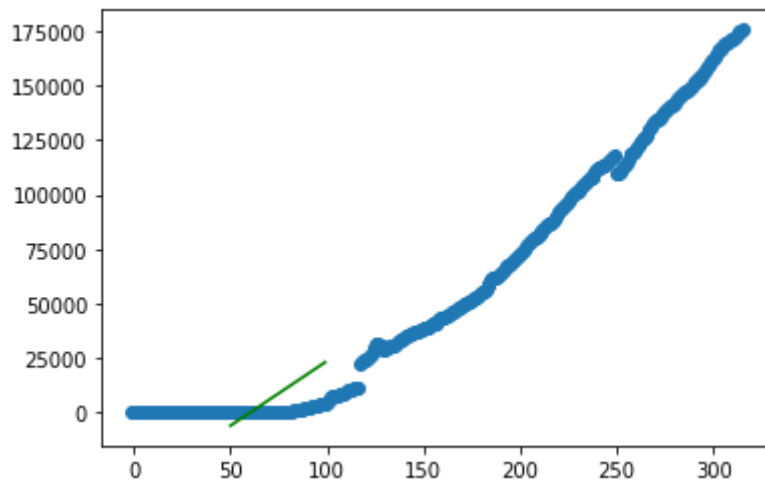
```
In [8]: #Vamos a comprobar:
# Quiero predecir cuántos "Casos" voy a obtener por en el día 100,
# según nuestro modelo, hacemos:
y_prediccion = regr.predict([[100]])
print('Total de casos', int(y_prediccion))
```

```
Total de casos 23672
```

```
In [9]: #Graficar
plt.scatter(x, y)
x_real = np.array(range(50, 100))
```

```
print(x_real)
plt.plot(x_real, regr.predict(x_real.reshape(-1, 1)), color='green')
plt.show()
```

```
[50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73
 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97
 98 99]
```



## El modelo logístico

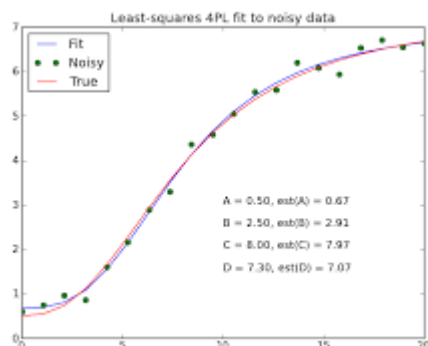
El modelo logístico se ha utilizado ampliamente para describir el crecimiento de una población. Una infección puede describirse como el crecimiento de la población de un agente patógeno, por lo que un modelo logístico parece razonable. La expresión más genérica de una función logística es:

$$f(x, a, b, c) = \frac{c}{1 + e^{-(x-b)/a}}$$

En esta fórmula, tenemos la variable  $x$  que es el tiempo y tres parámetros:  $a$ ,  $b$ ,  $c$ .

- $a$  se refiere a la velocidad de infección
- $b$  es el día en que ocurrieron las infecciones máximas
- $c$  es el número total de personas infectadas registradas al final de la infección

A continuación se puede apreciar un ejemplo de regresión logística



Definamos la función en Python y realicemos el procedimiento de ajuste de curva utilizado para el crecimiento logístico.

```
In [10]: def modelo_logistico(x,a,b):
          res= a+b*np.log(x)
```



```
return res
```

```
exp_fit = curve_fit(modelo_logistico,x,y) #Extraemos los valores de los parametros
print(exp_fit)
```

```
(array([1., 1.]), array([[inf, inf],
                        [inf, inf]]))
```

C:\Users\Jessica\anaconda3\lib\site-packages\ipykernel\_launcher.py:2: RuntimeWarning: divide by zero encountered in log

C:\Users\Jessica\anaconda3\lib\site-packages\ipykernel\_launcher.py:2: RuntimeWarning: invalid value encountered in log

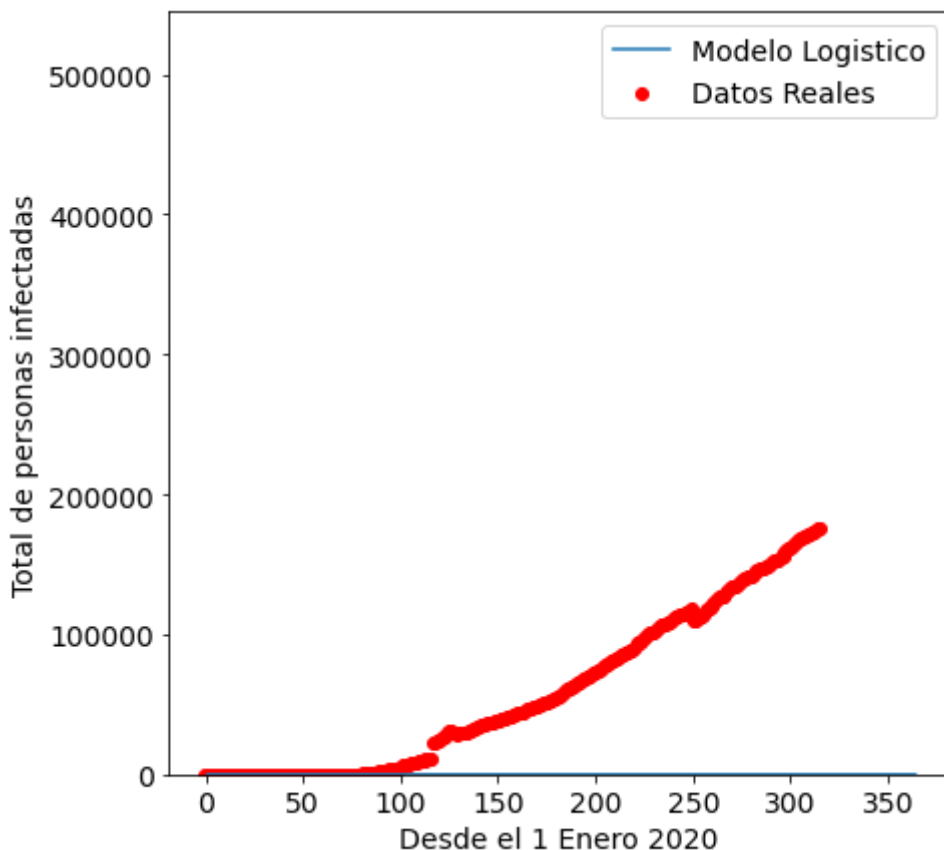
C:\Users\Jessica\anaconda3\lib\site-packages\scipy\optimize\minpack.py:808: OptimizeWarning: Covariance of the parameters could not be estimated  
category=OptimizeWarning)

## Graficas

```
In [11]: pred_x = list(range(min(x),max(x)+50)) # Predecir 50 dias mas
plt.rcParams['figure.figsize'] = [7, 7]
plt.rc('font', size=14)
# Real data
plt.scatter(x,y,label="Datos Reales",color="red")
# Predicted exponential curve
plt.plot(pred_x, [modelo_logistico(i,exp_fit[0][0],exp_fit[0][1]) for i in pred_x],
plt.legend()
plt.xlabel("Desde el 1 Enero 2020")
plt.ylabel("Total de personas infectadas")
plt.ylim((min(y)*0.9,max(y)*3.1)) # Definir los limites de Y
plt.show()
```

C:\Users\Jessica\anaconda3\lib\site-packages\ipykernel\_launcher.py:2: RuntimeWarning: invalid value encountered in log

C:\Users\Jessica\anaconda3\lib\site-packages\ipykernel\_launcher.py:2: RuntimeWarning: divide by zero encountered in log



## Modelo exponencial

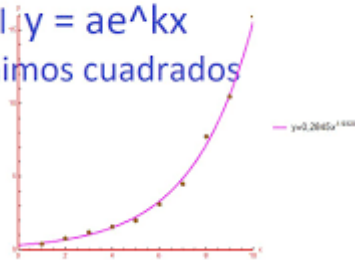
Mientras que el modelo logístico describe un crecimiento de infección que se detendrá en el futuro, el modelo exponencial describe un crecimiento de infección imparable. Por ejemplo, si un paciente infecta a 2 pacientes por día, después de 1 día tendremos 2 infecciones, 4 después

$$f(x, a, b, c) = a \cdot e^{b(x-c)}$$

de 2 días, 8 después de 3 y así sucesivamente.

A continuación se tiene un ejemplo de regresión exponencial

Curva de ajuste para una función tipo  
exponencial  $y = ae^{kx}$   
usando mínimos cuadrados



In [1]: `# Implementar`

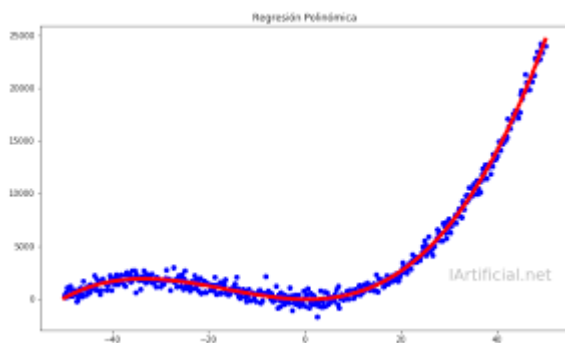
## Modelo polinomial

Predicción de una variable de respuesta cuantitativa a partir de una variable predictora cuantitativa, donde la relación se modela como una función polinomial de orden n (esto significa que pueden tener de diferentes exponentes o grados y se debe ir probando)

Se puede tener una ecuación con diferentes grados

$$y = a_0 + a_1x + a_2x^2 + a_3x^3 + \dots + a_nx^n + \epsilon$$

Ejemplo de una regresión polinómica de grado 4.



In [12]: `# Implementar`

```
# Se puede implementar modelos adicionales, en caso de ser así explicar o dar una in
# Se tomara como puntos adicionales al trabajo.
from sklearn.linear_model import LinearRegression
from sklearn.preprocessing import PolynomialFeatures

pf = PolynomialFeatures(degree = 4)
X = pf.fit_transform(np.array(x).reshape(-1, 1))
```

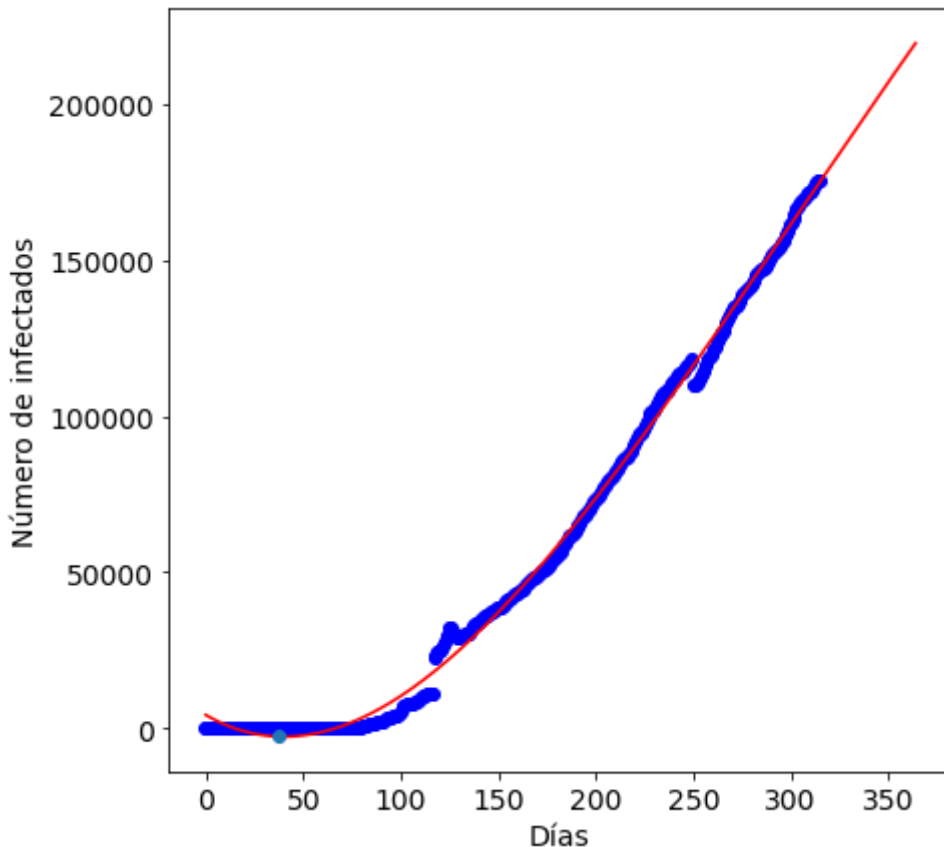
```

regresion_lineal = LinearRegression()
regresion_lineal.fit(X, y)
pred_x = list(range(0,max(x)+50))

puntos = pf.fit_transform(np.array(pred_x).reshape(-1, 1))
prediccion_entrenamiento = regresion_lineal.predict(puntos)
respuesta = round(-(prediccion_entrenamiento[37]))
print ('En los próximo 7 días habran :',respuesta, 'contagiados')
plt.plot(pred_x, prediccion_entrenamiento, color='red')
plt.scatter(x,y,label="Datos Reales",color="blue")
plt.xlabel("Días")
plt.ylabel("Número de infectados")
plt.plot(37,prediccion_entrenamiento[37], 'o')
plt.show()

```

En los próximo 7 días habran : 2802.0 contagiados



## Analisis

Mediante el método de regresión lineal podemos predecir el número de infectados por covid-19 en Ecuador mediante datos encontrados en internet, este método nos ayuda a tener una aproximación para de esta manera poder observar cual es el comportamiento de la pandemia en el futuro.

## Conclusiones

Como conclusión tenemos que la regresión Polinomial es un modelo matemático que nos ayuda a predecir el futuro en este caso sobre la pandemia del covid-19, este modelo es subjetivo por lo que no siempre acertara con veracidad lo que sucederá a futuro, pero si nos dará una idea de lo que sucederá y de esta manera las personas puedan tomar consciencia y prevenir el virus.

## Criterio personal (politico, economico y social de la situacion)

En el ámbito político las autoridades de nuestro país deben regir medidas de seguridad, de salud y realizar planificaciones que ayuden al bienestar de la sociedad para de esa manera reducir el contagio y evitar mas muertes, en el ámbito económico esta pandemia a afectado a muchos negocios que algunos tuvieron que cerrar, existe una tasa alta de desempleo por lo que las personas ya no tienen dinero y puede ocasionar enfermedades como la desnutrición y otras, pero también existen personas que han emprendido un negocio vendiendo objetos mediante redes sociales, y en el ámbito social puedo decir que esta pandemia a afectado a todas las personas de nuestro país como al mundo entero, tanto en la salud, ingresos y empleo.

## Referencias

- [https://www.researchgate.net/publication/340092755\\_Infeccion\\_del\\_Covid-19\\_en\\_Colombia\\_Una\\_comparacion\\_de\\_modelos\\_logisticos\\_y\\_exponenciales\\_aplicados\\_a\\_la\\_infeccion](https://www.researchgate.net/publication/340092755_Infeccion_del_Covid-19_en_Colombia_Una_comparacion_de_modelos_logisticos_y_exponenciales_aplicados_a_la_infeccion)
- <https://www.aprendemachinelearning.com/regresion-lineal-en-espanol-con-python/>

