UNIVERSIDAD POLITECNICA SALESIANA

Alumno: Juan Cañar.

Docente: Ing. Diego Quisi

Covid-19 infección en Ecuador. Modelos matemáticos y predicciones

Una comparación de modelos, lineal, polilnomico,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 crea 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 [2]:

```
# Importar las librerias para el analasis
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
from xml.dom import minidom
```

In [3]:

```
# Actualizar los datos (URL)
#datos tomados de https://www.google.com/covid19/mobility/
#https://github.com/owid/covid-19-data/tree/master/public/data
#url = 'Libro1.xlsx'
url = 'owid-covid-data.csv'
df = pd.read_csv(url)
df= df.fillna(1)
df
```

Out[3]:

	iso_code	continent	location	date	total_cases	new_cases	new_cases_smoothed	total_
0	ECU	South America	Ecuador	2019-12-31	1.0	0.0	1.000	
1	ECU	South America	Ecuador	2020-01-01	1.0	0.0	1.000	
2	ECU	South America	Ecuador	2020-01-02	1.0	0.0	1.000	
3	ECU	South America	Ecuador	2020-01-03	1.0	0.0	1.000	
4	ECU	South America	Ecuador	2020-01-04	1.0	0.0	1.000	
...	...	...	...	...	...	...	...	
305	ECU	South America	Ecuador	2020-10-31	167147.0	845.0	1268.143	1
306	ECU	South America	Ecuador	2020-11-01	168192.0	1045.0	1225.429	1
307	ECU	South America	Ecuador	2020-11-02	169194.0	1002.0	1079.857	1
308	ECU	South America	Ecuador	2020-11-03	169562.0	368.0	1054.857	1
309	ECU	South America	Ecuador	2020-11-04	170110.0	548.0	988.286	1

310 rows × 49 columns

Imprimos los resultados y agregamos el numero del dia

In [4]:

```
df = df[df['location'].isin(['Ecuador'])] #Filtro la Informacion solo para Ecuador
df = df.loc[:,['date','total_cases','iso_code']] #Selecciono las columnas de analasis
# 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("2019-12-30

df
#data[~data.isin([np.nan, np.inf, -np.inf]).any(1)]
```

Out[4]:

	date	total_cases	iso_code
0	1	1.0	ECU
1	2	1.0	ECU
2	3	1.0	ECU
3	4	1.0	ECU
4	5	1.0	ECU
...	...	...	...
305	306	167147.0	ECU
306	307	168192.0	ECU
307	308	169194.0	ECU
308	309	169562.0	ECU
309	310	170110.0	ECU

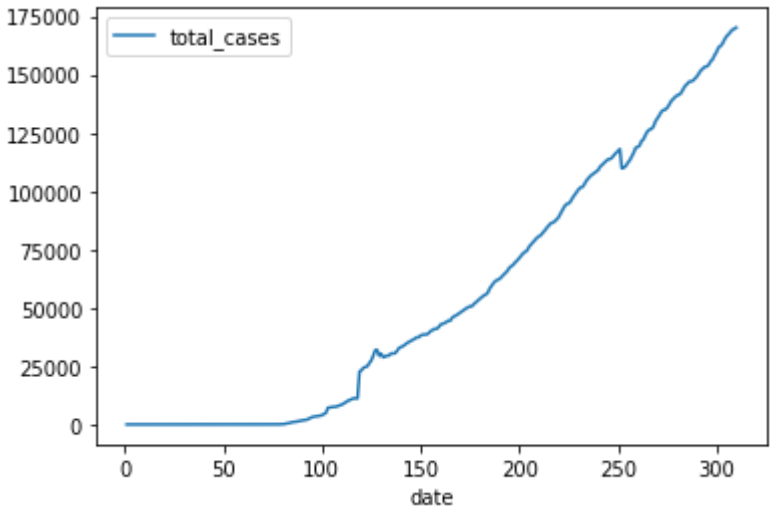
310 rows × 3 columns

In [5]:

```
 df.plot(x ='date', y='total_cases')
#df[~df.isin([np.nan, np.inf, -np.inf]).any(1)]
#df['Ecuador']=df['Ecuador'].fillna(0) #Reemplaza con ceros
#df = df.replace(np.nan, 1)
#df.describe()
```

Out[5]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x20d95904a48>
```



Ahora podemos analizar los cuatro modelos que tomaré en el examen, que son la función lineal, polinomica,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

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 [6]:

```
x = list(df.iloc [:, 0]) # Fecha
y = list(df.iloc [:, 1]) # Numero de casos
# Creamos el objeto de Regresión Lineal
regr = linear_model.LinearRegression()
print(x)
# Entrenamos nuestro modelo
regr.fit(np.array(x).reshape(-1, 1) ,y)

# Veamos los coeficienetes 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
```

```
[1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21,
22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40,
41, 42, 43, 44, 45, 46, 47, 48, 49, 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, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 11
3, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 12
8, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 14
3, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153, 154, 155, 156, 157, 15
8, 159, 160, 161, 162, 163, 164, 165, 166, 167, 168, 169, 170, 171, 172, 17
3, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183, 184, 185, 186, 187, 18
8, 189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 199, 200, 201, 202, 20
3, 204, 205, 206, 207, 208, 209, 210, 211, 212, 213, 214, 215, 216, 217, 21
8, 219, 220, 221, 222, 223, 224, 225, 226, 227, 228, 229, 230, 231, 232, 23
3, 234, 235, 236, 237, 238, 239, 240, 241, 242, 243, 244, 245, 246, 247, 24
8, 249, 250, 251, 252, 253, 254, 255, 256, 257, 258, 259, 260, 261, 262, 26
3, 264, 265, 266, 267, 268, 269, 270, 271, 272, 273, 274, 275, 276, 277, 27
8, 279, 280, 281, 282, 283, 284, 285, 286, 287, 288, 289, 290, 291, 292, 29
3, 294, 295, 296, 297, 298, 299, 300, 301, 302, 303, 304, 305, 306, 307, 30
8, 309, 310]
Coefficients:
 [582.45641984]
Independent term:
 -35705.066833698686
```

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

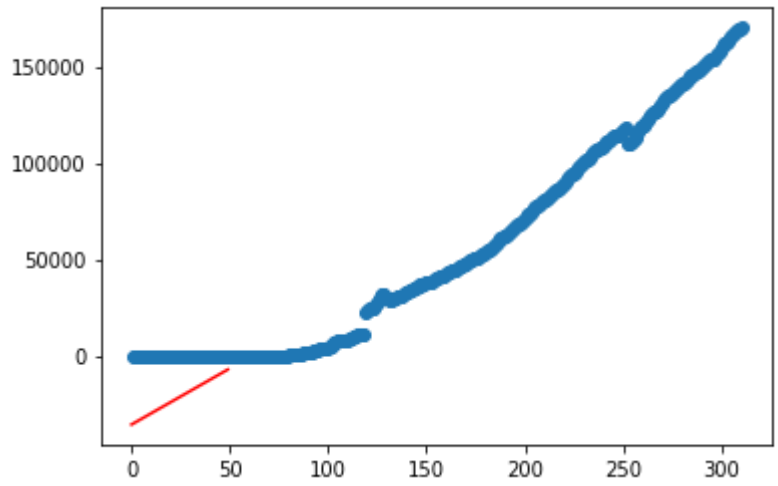
In [9]:

```
#Vamos a comprobar:
# Quiero predecir cuántos "Casos" voy a obtener por en el dia 100,
# según nuestro modelo, hacemos:
cantidad = int(input("ESCRIBA NUMERO DE DIAS: "))
y_prediccion = regr.predict([[cantidad]])
print('Para el dia',cantidad ,'hay estos casos: ',int(y_prediccion))
```

```
ESCRIBA NUMERO DE DIAS: 20
Para el dia 20 hay estos casos:  -24055
```

In [10]:

```
#Graficar
plt.scatter(x, y)
x_real = np.array(range(0, 50))
#print(x_real)
plt.plot(x_real, regr.predict(x_real.reshape(-1, 1)), color='red')
plt.show()
```



El modelo logistico

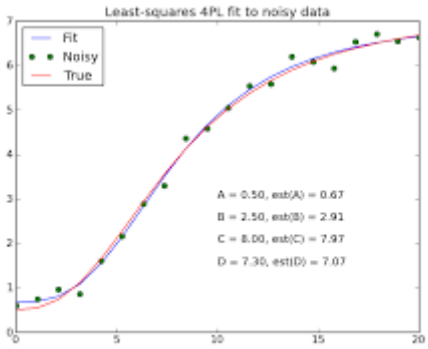
El modelo logístico se ha utilizado ampliamente para describir el crecimiento de una población. Una infección puede describirse como el crecimiento de la población de un agente patógeno, por lo que un modelo logístico parece razonable . La expresión más genérica de una función logística es:

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

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 continuacion se puede apreciar un ejemplo de regresion logistica



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

In [11]:

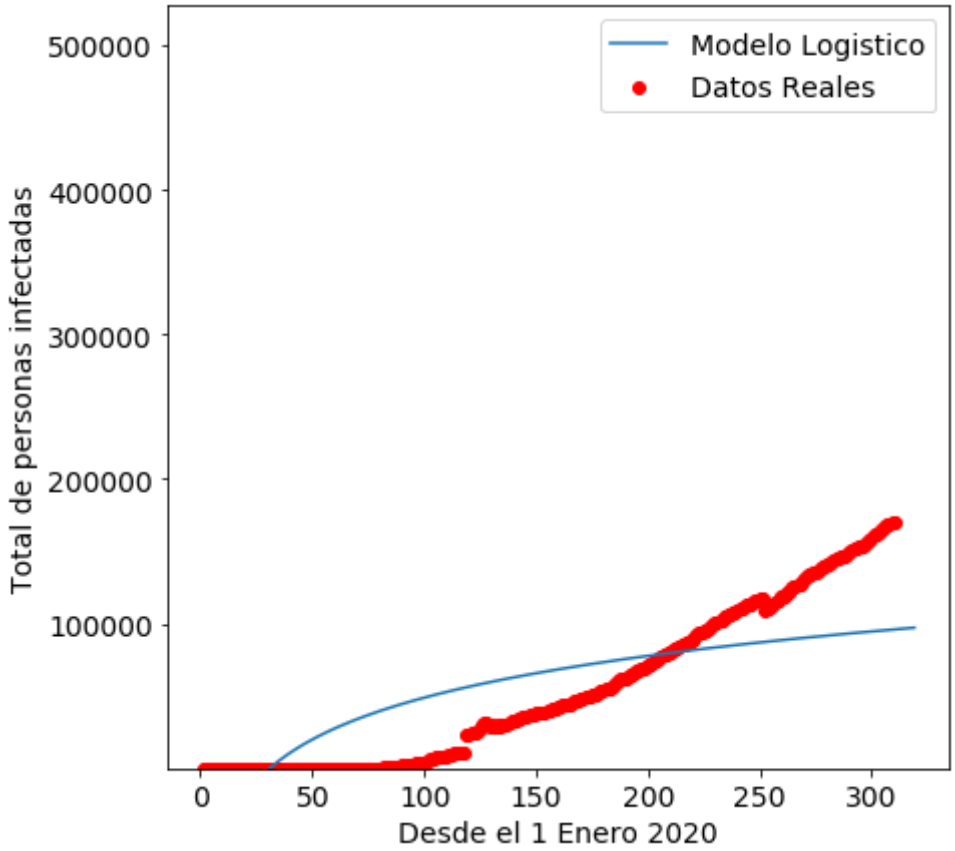
```
def modelo_logistico(x,a,b):
    res= a+b*np.log(x)
    return res

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

```
(array([-143165.45891415,   41701.64700227]), array([[ 1.07640081e+08, -2.17
684167e+07],
       [-2.17684167e+07,  4.58399227e+06]]))
```

Graficas

In [12]:

```
pred_x = list(range(min(x),max(x)+10)) # 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], label="
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()
```

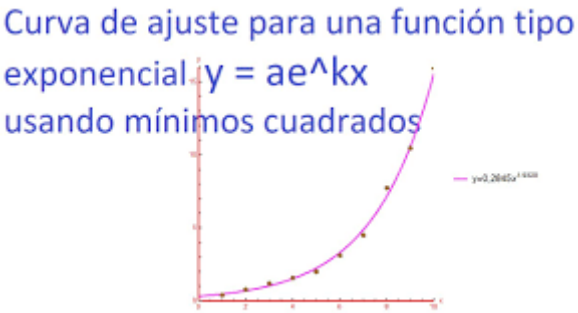


Modelo exponencial

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

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

A continuacion se tiene un ejemplo de regresion exponencial



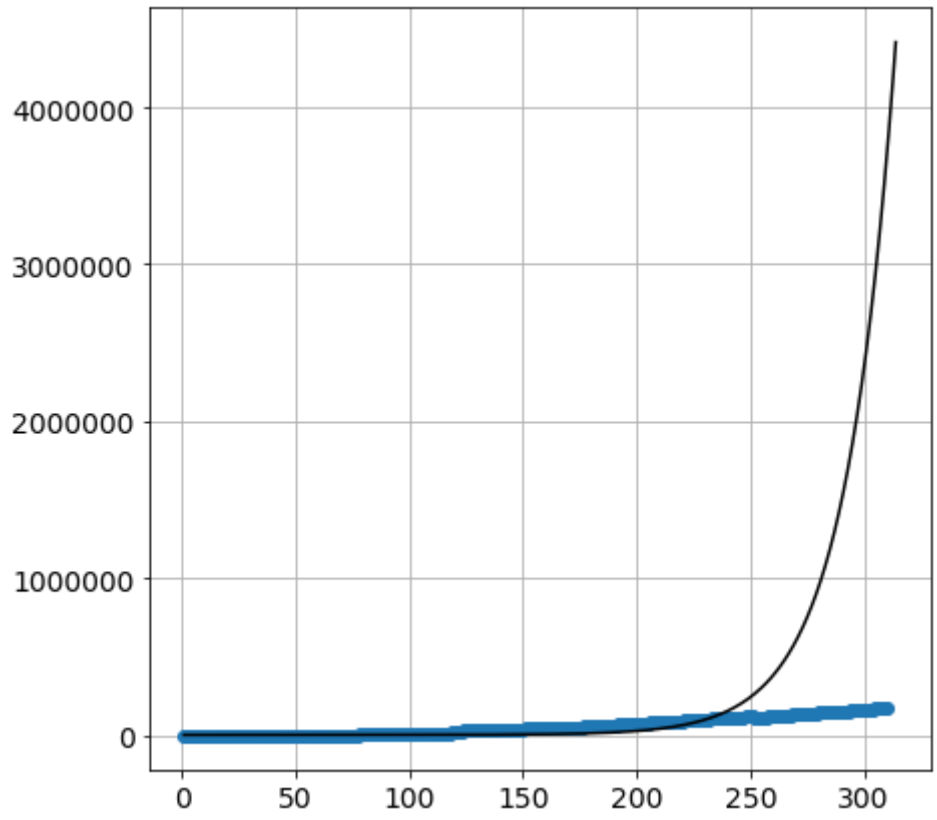
In [13]:

```
curve_fit = np.polyfit((x),np.log(y), deg=1)
print(curve_fit)
```

```
[0.04550896 1.01078624]
```

In [14]:

```
pred_x = np.array(list(range(min(x),max(x)+5)))
yx = np.exp(curve_fit[1]) * np.exp(curve_fit[0]*pred_x)
plt.plot(x,y,'o')
plt.plot(pred_x,yx,color="black")
plt.grid(True)
```



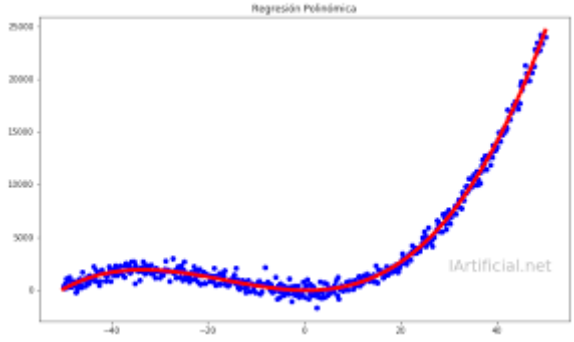
Modelo polinomial

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 exponenciales o grados y se debe ir probando)

Se puede tener una ecuacion con diferentes grados

y = a0 + a1x + a2x^2 + a3x^3 + … + anx^n + ε

Ejemplo de una regresion polinomica de grado 4.



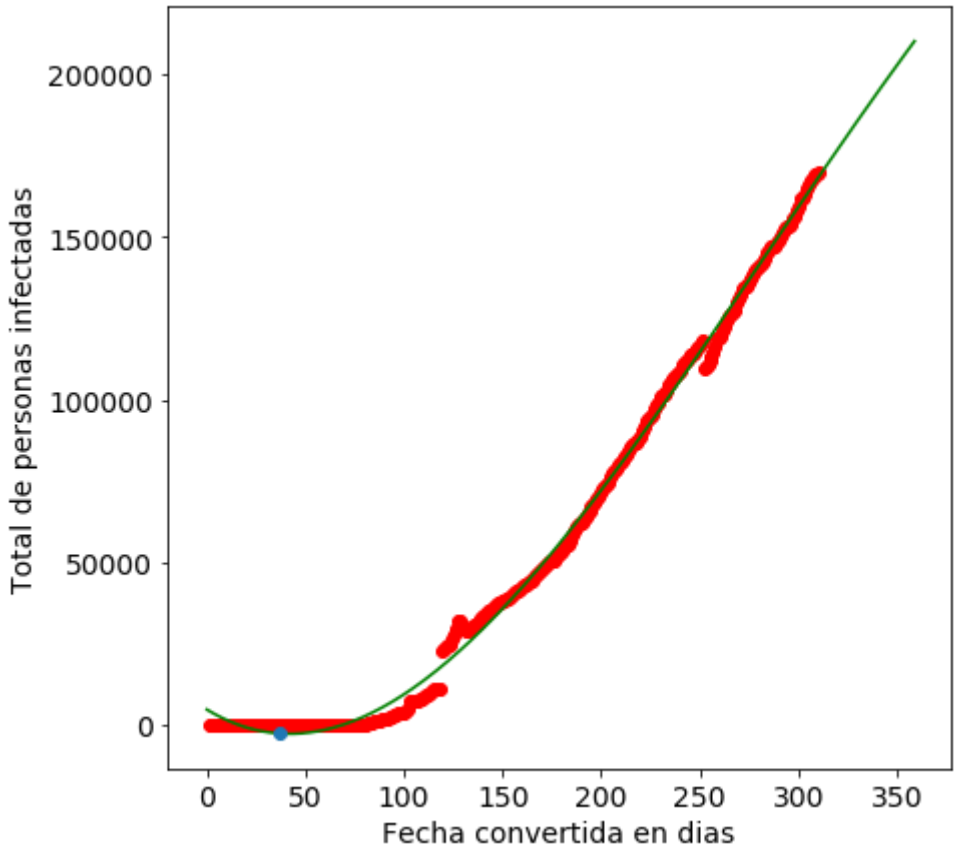
Deber:

Implementar el modelo Polinomial

In [16]:

```
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('----------------RESPUESTA OBTENIDA en 7 DIAS------------')
print ('Se hace cuenta desde el dia(30) y se suma 7 dias, PREDICCION:',respuesta, 'contagia
plt.plot(pred_x, prediccion_entrenamiento, color='green')
plt.scatter(x,y,label="Datos Reales",color="red")
plt.xlabel("Fecha convertida en dias")
plt.ylabel("Total de personas infectadas")
plt.plot(37,prediccion_entrenamiento[37], 'o')
plt.show()
```

```
----------------RESPUESTA OBTENIDA en 7 DIAS------------
Se hace cuenta desde el dia(30) y se suma 7 dias, PREDICCION: 2668.0 contagi
ados
```



Conclusiones

Como conclusion se puede decir que estos modelos matematicos son herramientas que nos ayudan a simular o predecir en un futuro los eventos que pasaran, aunque cabe recalcar que estos modelos son bastante subjetivos, lo que podria no acertar en su totalidad los eventos, pero si nos podria dar una referencia intermedia para poder tener un concepto claro, de como va ha afectarnos o que probabilidades existen de incremento o decremento.

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

El ambito politico frente al covid juega un papel muy importante, ya que un buen manejo de los recursos del pais, puede dar una pronta respuesta a la pandemia mediante un buen manejo de entidades de salud y una buena planificacion, esto de cierta forma ayudaria a reducir los contagios y a manejar de mejor manera, ya que en pleno rebrote fuimos uno de los paises latinoamericanos mas afectado y con un alto indice de mortalidad como consecuencia del mal manejo y de la descarada corrupcion que existe, en cuanto a lo economico, es una etapa bastante dificil para el pais y su comunidad, debido a que la tasa de subempleo incrementa y tambien existen reduccion en la nomida de empleados, lo que trae consigo como consecuencia, que las personas no tengan dinero para mantenerse, tampoco puden acceder a la canasta basica, que desencadena en varios factores tales como desnutricion, enfermedades, personas sin hogar por la falta de posibilidades economicas, otros buscan obtener ingresos mediante microemprendimientos, para aliviar de cierta forma este problema, y en lo social haciendo referencia a puntos mencionados anteriormente se puede decir que afecta a todas las personas, en distintos ambitos tales como empleo, ingresos, comercios, salud, vivieda.

Referencias

- https://www.researchgate.net/publication/340092755_Infeccion_del_Covid-19_en_Colombia_Una_comparacion_de_modelos_logisticos_y_exponenciales_aplicados_a_la_infeccion_pc (https://www.researchgate.net/publication/340092755_Infeccion_del_Covid-19_en_Colombia_Una_comparacion_de_modelos_logisticos_y_exponenciales_aplicados_a_la_infeccion_pc
- https://www.aprendemachinelearning.com/regresion-lineal-en-espanol-con-python/ (https://www.aprendemachinelearning.com/regresion-lineal-en-espanol-con-python/)
- https://ourworldindata.org/coronavirus-data (https://ourworldindata.org/coronavirus-data)
- https://github.com/owid/covid-19-data/blob/master/public/data/owid-covid-codebook.csv (https://github.com/owid/covid-19-data/blob/master/public/data/owid-covid-codebook.csv)