



**Nombre:**

Tatiana Doménica Cardenas Jho

**Tema:**

Regresión de los casos de COVID-19 en Ecuador

**Asignatura:**

Simulación

**Docente:**

Ing. Diego Fernando Quisi Peralta

**Fecha:**

Cuenca, 25 de abril de 2021

## Enunciado:

Generar un modelo de regresión de los casos confirmados de COVID dentro del Ecuador, el mismo que permita predecir el comportamiento y/o predicción de la pandemia, tomar los datos desde el inicio e identificar etapas: Confinamiento, Toques de Queda, Feriados, etc.

## Desarrollo:

La presente práctica fue desarrollada mediante el modelo de regresión lineal y el modelo de regresión polinomial para poder predecir el número de contagios de COVID-19 en el Ecuador.

A continuación, se presenta los datos del dataset que fue utilizado para el desarrollo de la practica:

	iso_code	continent	location	date	total_cases	new_cases	new_cases_smoothed	total_deaths	new_deaths	new_deaths_smoothed	...	gdp_per_capita
38	ECU	South America	Ecuador	2020-04-08	4450	703	243.143	242.0	51.0	21.286	...	10581.936
386	ECU	South America	Ecuador	2021-03-22	312851	253	1479.000	16478.0	27.0	34.000	...	10581.936
133	ECU	South America	Ecuador	2020-07-12	67870	661	844.571	5047.0	16.0	38.000	...	10581.936
5	ECU	South America	Ecuador	2020-03-06	13	0	1.857	NaN	NaN	0.000	...	10581.936
179	ECU	South America	Ecuador	2020-08-27	111219	670	815.857	6471.0	61.0	38.714	...	10581.936

5 rows × 59 columns

Ilustración 1. Dataset COVID-19

	total_cases	new_cases	new_cases_smoothed	total_deaths	new_deaths	new_deaths_smoothed	total_cases_per_million	new_cases_per_million	new_deaths_per_million
count	407.000	407.000	402.000	394.000	394.000	402.000	407.000	407.000	407.000
mean	137704.005	852.130	849.336	9215.898	43.891	42.604	7804.996	48.298	48.298
std	100545.657	967.507	468.354	5646.413	196.902	75.717	5698.879	54.838	54.838
min	6.000	-7953.000	-525.000	2.000	0.000	0.000	0.340	-450.772	-450.772
25%	44440.000	353.000	562.572	4172.750	12.000	21.179	2518.837	20.008	20.008
50%	126419.000	765.000	869.143	11276.000	27.000	31.143	7165.367	43.360	43.360
75%	212012.000	1202.000	1123.750	14059.000	43.000	39.964	12016.736	68.128	68.128
max	346817.000	11536.000	2038.429	17293.000	3852.000	597.000	19657.418	653.855	653.855

8 rows × 54 columns

Ilustración 2. Descripción del Dataset COVID-19

## Pruebas:

### Regresión lineal:

La regresión lineal es un campo de estudio que enfatiza la relación estadística entre dos variables continuas conocidas como variables de predicción y respuesta.

1.

```
#PRUEBA 1

m = linear_regressor.coef_[0][0]
c = linear_regressor.intercept_[0]
#Prediccion de x dias desde la fecha inicial del Dataset 2020-03-01
dias=700
label1 = m*dias-c
print("Numero de casos a los ",dias," de la fecha inicial del dataset, se obtien
```

Numero de casos a los 700 de la fecha inicial del dataset, se obtiene una predicción de : 679202.5680659452

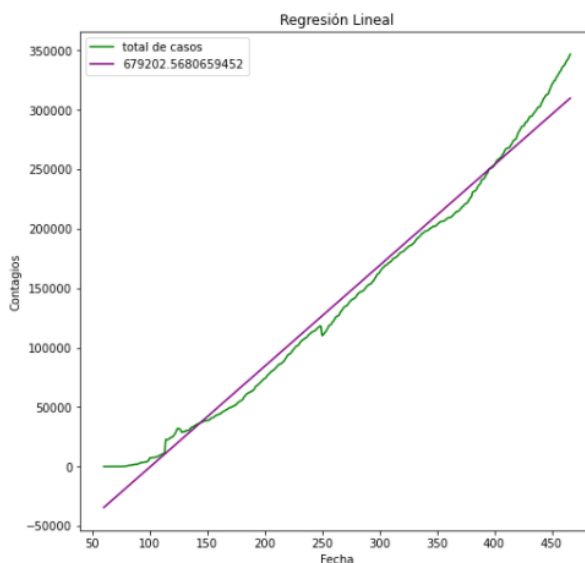


Ilustración 3. Prueba 1 de regresión lineal

2.

```
: #PRUEBA 2

m = linear_regressor.coef_[0][0]
c = linear_regressor.intercept_[0]
#Prediccion de x dias desde la fecha inicial del Dataset 2020-03-01
dias=750
label2 = m*dias-c
print("Numero de casos a los ",dias," de la fecha inicial del dataset, se obtien
```

Numero de casos a los 750 de la fecha inicial del dataset, se obtiene una predicción de : 721617.2395602312

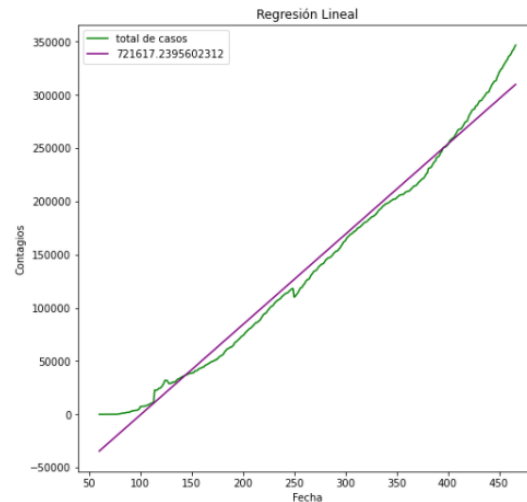


Ilustración 4. Prueba 2 de regresión lineal

3.

```
#PRUEBA 3

m = linear_regressor.coef_[0][0]
c = linear_regressor.intercept_[0]
#Prediccion de x dias desde la fecha inicial del Dataset 2020-03-01
dias=800
label3 = m*dias-c
print("Numero de casos a los ",dias," de la fecha inicial del dataset, se obtien
```

Numero de casos a los 800 de la fecha inicial del dataset, se obtiene una pre  
diccion de : 764031.9110545174

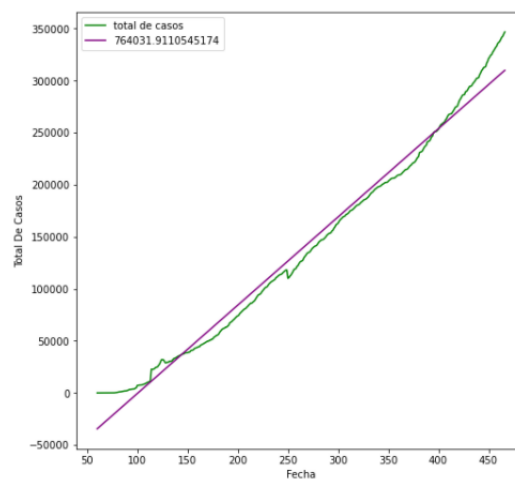


Ilustración 5. Prueba 3 de regresión lineal

4.

```
#PRUEBA 4

m = linear_regressor.coef_[0][0]
c = linear_regressor.intercept_[0]
#Prediccion de x dias desde la fecha inicial del Dataset 2020-03-01
dias=900
label4 = m*dias-c
print("Numero de casos a los ",dias," de la fecha inicial del dataset, se obtien
```

Numero de casos a los 900 de la fecha inicial del dataset, se obtiene una pre  
diccion de : 848861.2540430896

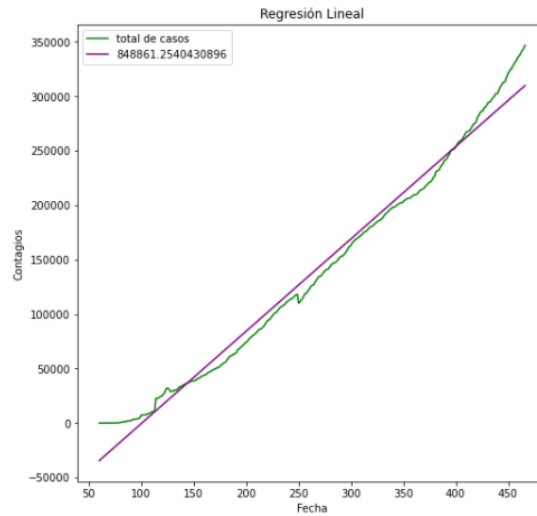


Ilustración 6. Prueba 4 de regresión lineal

5.

```
: #PRUEBA 5

m = linear_regressor.coef_[0][0]
c = linear_regressor.intercept_[0]
#Prediccion de x dias desde la fecha inicial del Dataset 2020-03-01
dias=968
label5 = m*dias-c
print("Numero de casos a los ",dias," de la fecha inicial del dataset, se obtien
```

Numero de casos a los 968 de la fecha inicial del dataset, se obtiene una pre  
diccion de : 906545.2072753186

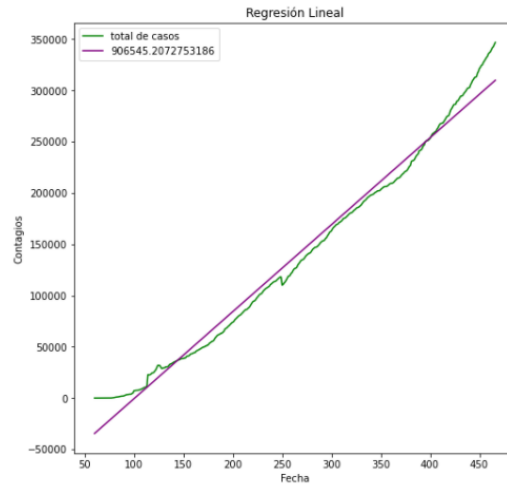


Ilustración 7. Prueba 5 de regresión lineal

### 📊 Regresión 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

```
#Prediccion de x dias desde la fecha inicial del Dataset 2020-03-01=dia 60
pol_prediccion = regresion_lineal.predict(pf.fit_transform([[700]]))
print("Prediccion para el dia 700 despues del dia 60:")
print(int(pol_prediccion))
```

Prediccion para el dia 700 despues del dia 60:  
5806171

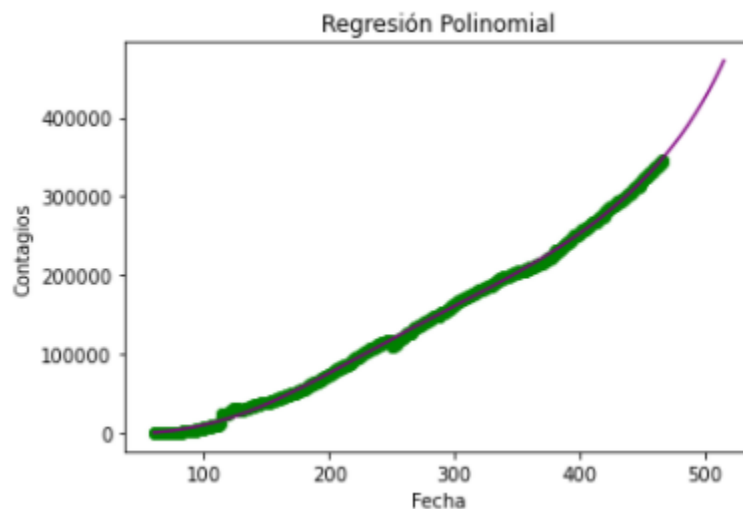


Ilustración 7. Prueba Regresión polinomial

### Regresión lineal vs Regresión polinomial:

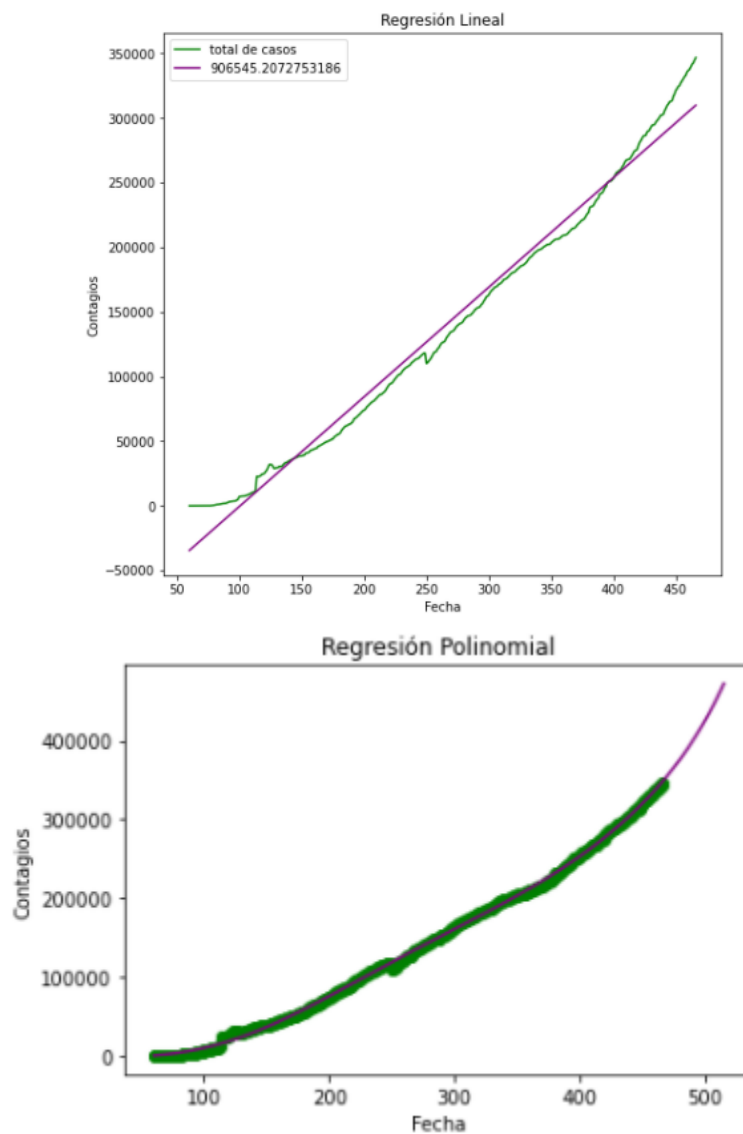


Ilustración 7. Regresión lineal vs regresión polinomial

Como se puede evidenciar en las gráficas presentes, la curva de contagios en la regresión polinomial se ajusta de mejor manera a comparación de la regresión lineal.

## **Conclusiones**

- La regresión polinomial consigue añadir curvatura al modelo introduciendo nuevos predictores que se obtienen al elevar todos o algunos de los predictores originales a distintas potencias, de tal manera que en la comparación presentada anteriormente se pudo observar como la regresión polinomial se ajustó de mejor manera a la curvatura de contagios.
- Los modelos lineales tienen la ventaja de ser fácilmente interpretables, sin embargo, pueden tener limitaciones importantes en capacidad predictiva.