

Ejercicio 4

```
In [19]: 1 import pandas as pd
          2 from sklearn.linear_model import LinearRegression
          3 import seaborn as sns
          4 import matplotlib.pyplot as plt
          5
          6 datos1 = pd.read_csv('datos(1).csv')
          7
          8 datos1
```

```
Out[19]:
```

	year	growth	vote	inc_party_candidate	other_candidate
0	1952	2.40	44.60	Stevenson	Eisenhower
1	1956	2.89	57.76	Eisenhower	Stevenson
2	1960	0.85	49.91	Nixon	Kennedy
3	1964	4.21	61.34	Johnson	Goldwater
4	1968	3.02	49.60	Humphrey	Nixon
5	1972	3.62	61.79	Nixon	McGovern
6	1976	1.08	48.95	Ford	Carter
7	1980	-0.39	44.70	Carter	Reagan
8	1984	3.86	59.17	Reagan	Mondale
9	1988	2.27	53.94	Bush, Sr.	Dukakis
10	1992	0.38	46.55	Bush, Sr.	Clinton
11	1996	1.04	54.74	Clinton	Dole

```

In [22]: 1 # Función para inferir el partido basándose en el nombre del candid
2 def infer_party(candidate_name):
3     if 'Bush' in candidate_name:
4         return 'Republican'
5     elif 'Reagan' in candidate_name:
6         return 'Republican'
7     elif 'Nixon' in candidate_name:
8         return 'Republican'
9     elif 'Eisenhower' in candidate_name:
10        return 'Republican'
11    elif 'Ford' in candidate_name:
12        return 'Republican'
13    elif 'McCain' in candidate_name:
14        return 'Republican'
15    else:
16        return 'Democrat'
17
18 # Crear columnas para votos de Demócratas y Republicanos
19 datos1['party'] = datos1['inc_party_candidate'].apply(infer_party)
20 datos1['democrat_votes'] = datos1.apply(lambda row: row['vote'] if
21 datos1['republican_votes'] = datos1.apply(lambda row: row['vote'] if
22
23
24 datos1

```

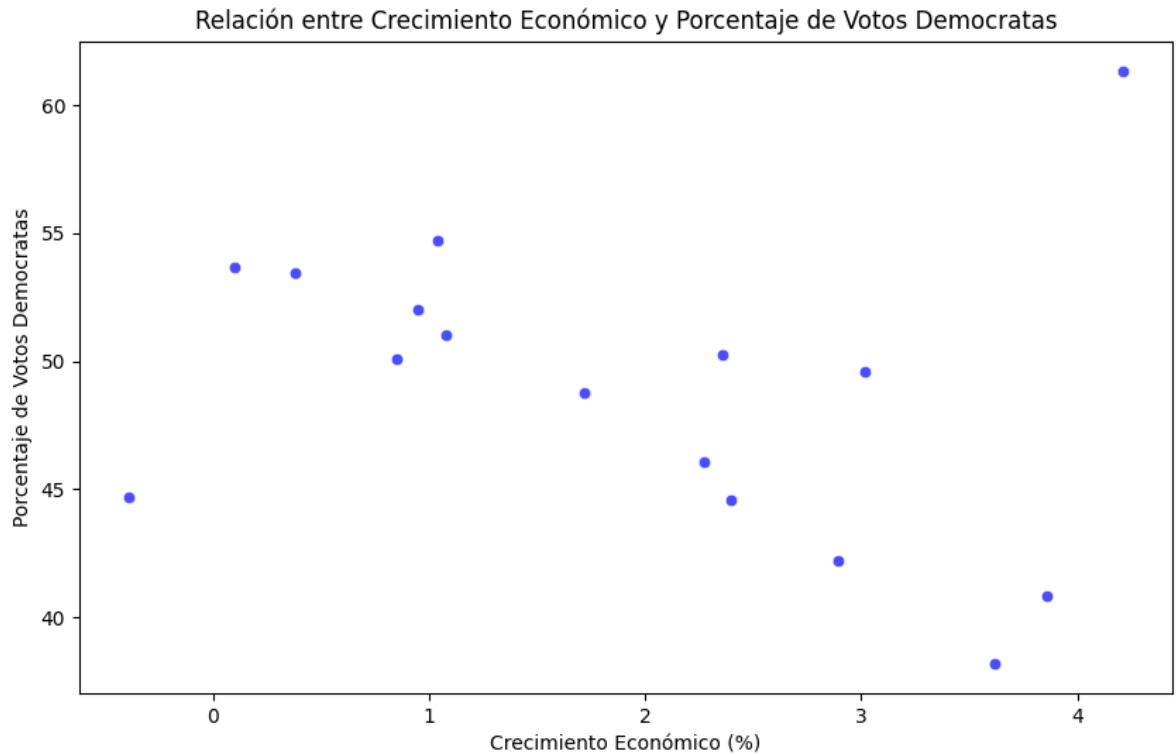
```

Out[22]:

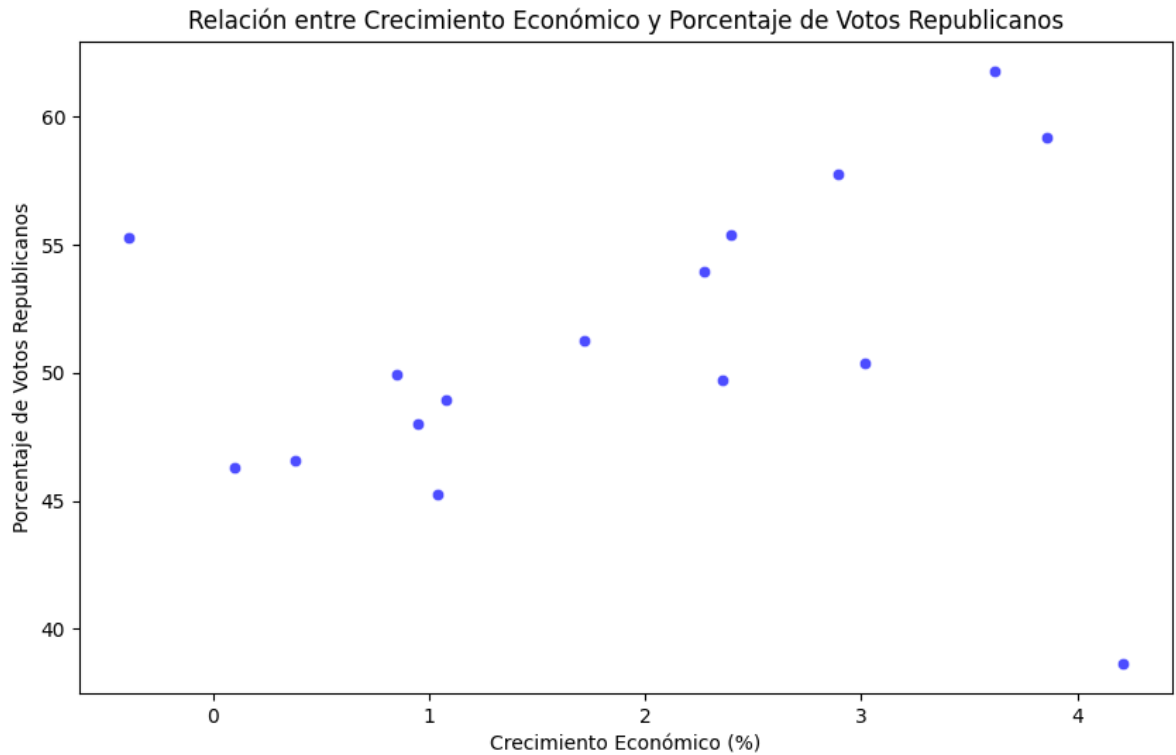
```

	year	growth	vote	inc_party_candidate	other_candidate	party	party_votes	democra
0	1952	2.40	44.60	Stevenson	Eisenhower	Democrat	44.60	
1	1956	2.89	57.76	Eisenhower	Stevenson	Republican	42.24	
2	1960	0.85	49.91	Nixon	Kennedy	Republican	50.09	
3	1964	4.21	61.34	Johnson	Goldwater	Democrat	61.34	
4	1968	3.02	49.60	Humphrey	Nixon	Democrat	49.60	
5	1972	3.62	61.79	Nixon	McGovern	Republican	38.21	
6	1976	1.08	48.95	Ford	Carter	Republican	51.05	
7	1980	-0.39	44.70	Carter	Reagan	Democrat	44.70	
8	1984	3.86	59.17	Reagan	Mondale	Republican	40.83	
9	1988	2.27	53.94	Bush, Sr.	Dukakis	Republican	46.06	
10	1992	0.38	46.55	Bush, Sr.	Clinton	Republican	53.45	
11	1996	1.04	54.74	Clinton	Dole	Democrat	54.74	
12	2000	2.36	50.27	Gore	Bush, Jr.	Democrat	50.27	
13	2004	1.72	51.24	Bush, Jr.	Kerry	Republican	48.76	
14	2008	0.10	46.32	McCain	Obama	Republican	53.68	
15	2012	0.95	52.00	Obama	Romney	Democrat	52.00	

```
In [24]: 1 plt.figure(figsize=(10, 6))
2 sns.scatterplot(x='growth', y='democrat_votes', data=datos1, color=
3 plt.title('Relación entre Crecimiento Económico y Porcentaje de Vot
4 plt.xlabel('Crecimiento Económico (%)')
5 plt.ylabel('Porcentaje de Votos Demócratas')
6 plt.show()
```



```
In [26]: 1 plt.figure(figsize=(10, 6))
2 sns.scatterplot(x='growth', y='republican_votes', data=datos1, color='blue')
3 plt.title('Relación entre Crecimiento Económico y Porcentaje de Votos Republicanos')
4 plt.xlabel('Crecimiento Económico (%)')
5 plt.ylabel('Porcentaje de Votos Republicanos')
6 plt.show()
7
```



```
In [27]: 1 # Define las variables independientes (X) y dependientes (Y)
2 X = datos1['growth'].values.reshape(-1, 1)
3 Y = datos1['republican_votes'].values
4
5 # Crea el modelo de regresión lineal
6 modelo_regresion = LinearRegression()
7
8 # Ajusta el modelo a tus datos
9 modelo_regresion.fit(X, Y)
10
11 # Imprime los coeficientes estimados
12 print('Intersección (beta_0):', modelo_regresion.intercept_)
13 print('Pendiente (beta_1):', modelo_regresion.coef_[0])
14
15 # Predicción para las elecciones entre Hillary y Trump con un creci
16 crecimiento_2015 = 2
17 prediccion = modelo_regresion.predict([[crecimiento_2015]])
18 print('Predicción para 2% de crecimiento en 2015 para Trump:', pred
19 print('Predicción para 2% de crecimiento en 2015 para Hilary:', 100
20
```

Intersección (beta_0): 49.15425067350383

Pendiente (beta_1): 1.0511195396554263

Predicción para 2% de crecimiento en 2015 para Trump: 51.2564897528146
9

Predicción para 2% de crecimiento en 2015 para Hilary: 48.743510247185
31

Para desarrollar un modelo de regresión lineal, primero necesitamos analizar los datos y comprender la relación entre el porcentaje de votos para el partido del presidente saliente (Y) y el crecimiento económico medido en ingreso personal en el año previo a las elecciones (X).

La regresión lineal puede representarse como:

$$Y = \beta_0 + \beta_1 X + \epsilon$$

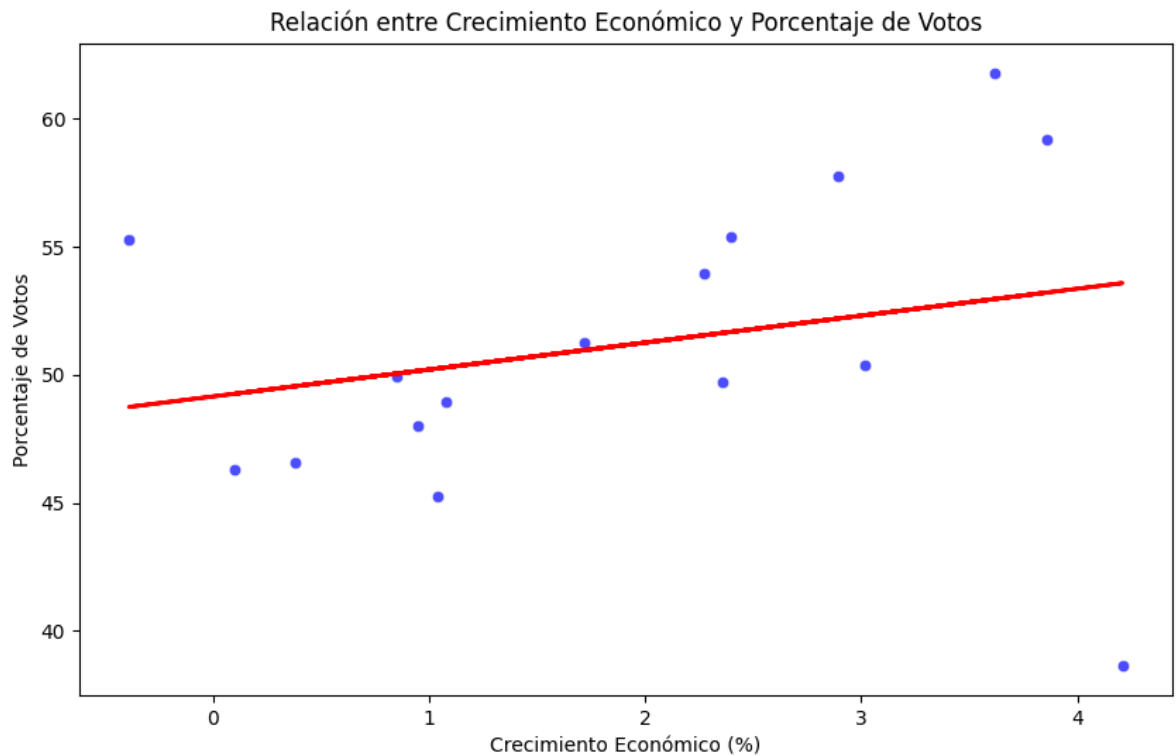
Donde:

- (Y) es el porcentaje de votos para el partido del presidente saliente.
- (X) es el porcentaje de crecimiento económico en el año previo a las elecciones.
- (β_0) es la intersección y (β_1) es la pendiente de la línea de regresión.
- (ϵ) es el término de error.

Utilizamos técnicas de regresión lineal para estimar los valores de β_0 y β_1 a partir de los datos.

In [28]:

```
1 # Predicciones
2 predicciones = modelo_regresion.predict(X)
3
4 # Gráfico de dispersión y línea de regresión
5 plt.figure(figsize=(10, 6))
6 sns.scatterplot(x='growth', y='republican_votes', data=datos1, color='blue')
7 plt.plot(datos1['growth'], predicciones, color='red', linewidth=2)
8 plt.title('Relación entre Crecimiento Económico y Porcentaje de Votos')
9 plt.xlabel('Crecimiento Económico (%)')
10 plt.ylabel('Porcentaje de Votos')
11 plt.show()
```



Comentarios

De acuerdo a los datos, nos percatamos de que los datos estan incompletos para hacer una predicción para un partido en especifico, ya que los votos que muestran son del partido actual en el poder, es decir, están mezclados entre partidos, por lo que hacer una predicción con estos datos sería incorrecto. Por lo que creamos dos columnas para tener los votos por partido politico.

Haciendo un analisis de los datos, notamos que se ajustan mejor a una regresión lineal los votos del partido republicano por lo que con estos datos predecimos los votos para Trump y los restantes para el partido demócrata Hillary

