

```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
import sqlalchemy
import warnings
import json
warnings.filterwarnings('ignore')
```

Carga de archivo de configuracion

```
with open("config.json", "r") as f:
    config = json.load(f)
```

Creacion de engine para conectar con la bd

```
engine =
sqlalchemy.create_engine("mariadb+mariadbconnector://root:Maisicual123
@127.0.0.1:3306/data_analysis_ss2")
```

Carga de datos desde la base de datos

```
# Carga de datos de las muertes por municipio
df_muertes_municipio = pd.read_sql(config["QUERY_MUERTES_MUNICIPIO"],
engine)

# Cargar datos de tabla general_data_by_fecha
df_general_data = pd.read_sql(config["QUERY_MUERTES_GENERALES"],
engine)
```

EDA Monovariable

Datos Cuantitativos

Variable: Nuevas muertes

```
df_nuevas_muertes = df_general_data['muertes_nuevas']
```

Describe

```
df_nuevas_muertes.describe()
```

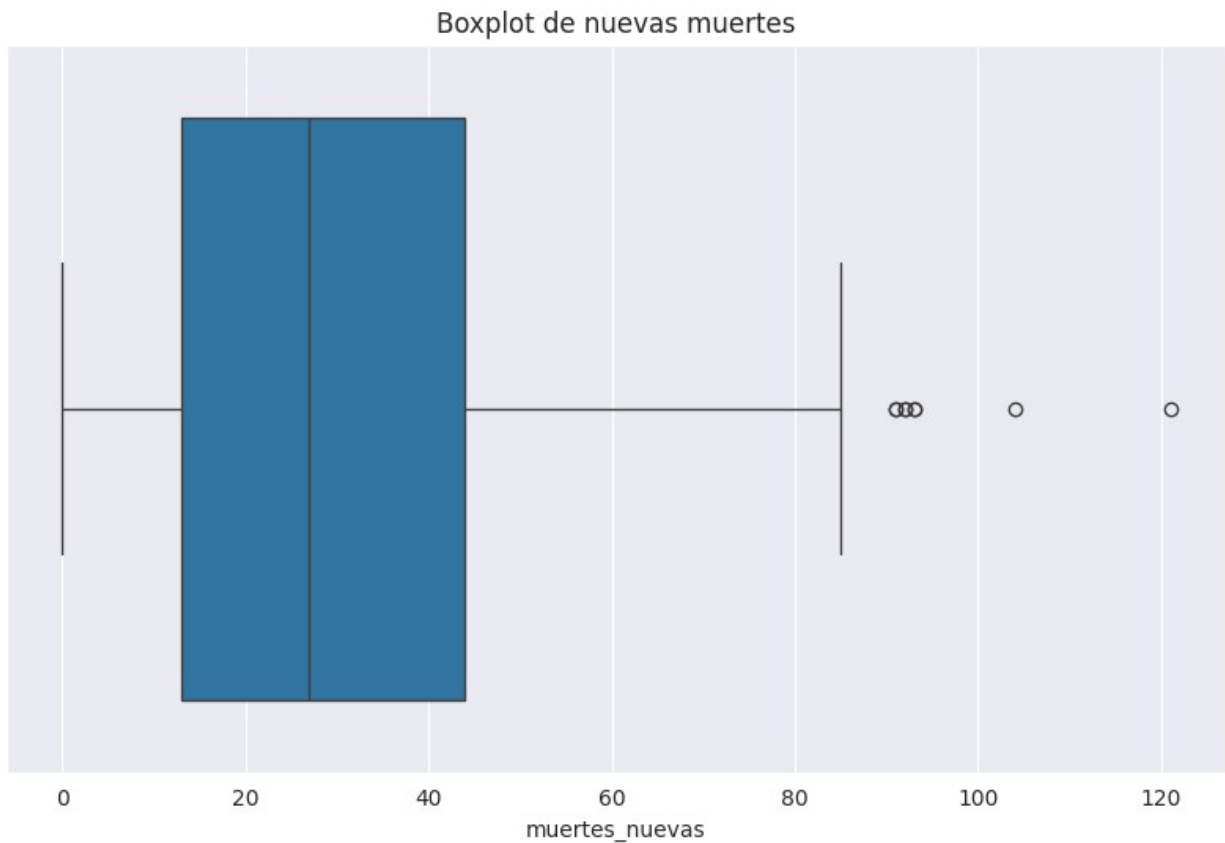
```
count    365.000000
mean     30.956164
std      22.509662
min       0.000000
25%      13.000000
50%      27.000000
75%      44.000000
max      121.000000
Name: muertes_nuevas, dtype: float64
```

Valores unicos

```
df_nuevas_muertes.unique()
array([ 10,  7,  6,  2, 24, 40, 29, 44, 27, 26,  0, 92,
        34,
        43, 35, 30, 46, 31, 36,  9,  4, 17, 39, 25, 49,
        38,
        33, 37,  8, 22, 12, 14, 19, 15, 32,  5, 20, 16,
        3,
        21, 18, 28, 13, 11, 23, 50, 41, 60, 66, 48, 42,
        58,
        68, 67, 69, 78, 64, 91, 74, 54, 70, 57, 53, 62,
        79,
        104, 63, 81, 73, 80, 76, 121, 45, 93, 56, 59, 72,
        61,
        75, 51, 52, 65, 85,  1])
```

Grafica de caja

```
plt.figure(figsize=(10, 6))
ax = sns.boxplot(x=df_nuevas_muertes)
ax.set_title('Boxplot de nuevas muertes')
plt.show()
```

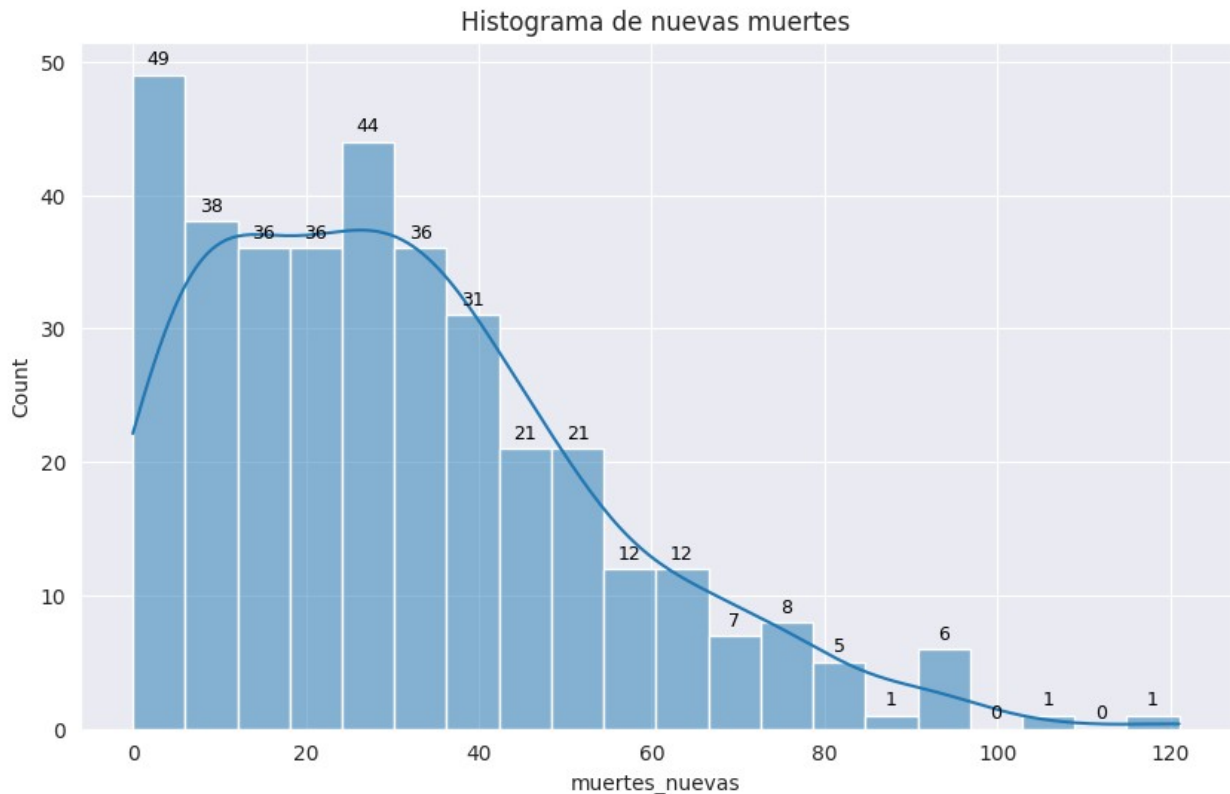


Histograma

```
# Histograma
plt.figure(figsize=(10, 6))
ax = sns.histplot(df_nuevas_muertes, kde=True, bins=20)

for p in ax.patches:
    ax.annotate(f'{p.get_height():.1f}', (p.get_x() + p.get_width() / 2.,
    p.get_height()),
    ha='center', va='baseline', fontsize=9, color='black',
    xytext=(0, 5),
    textcoords='offset points')

ax.set_title('Histograma de nuevas muertes')
plt.show()
```



Analisis sobre datos cuantitativos sobre nuevas muertes

En el diagrama de caja podemos observar unos datos que estan fuera del rango intercuartilico, estos datos son los que se encuentran por encima de 100 muertes, estos datos son atipicos, pero no se pueden eliminar ya que son datos reales. En el histograma podemos observar que la mayoría de los datos se encuentran entre 0 y 50 muertes, con una media de 30.95.

Variable: Muertes acumuladas

```
df_nuevas_muertes = df_general_data['muertes_acumuladas']
```

Describe

```
df_nuevas_muertes.describe()
```

```
count      365.000000
mean      10108.531507
std       3672.770919
min       4813.000000
25%       6860.000000
50%       9282.000000
75%      13564.000000
```

```
max      16102.000000
Name: muertes_acumuladas, dtype: float64
```

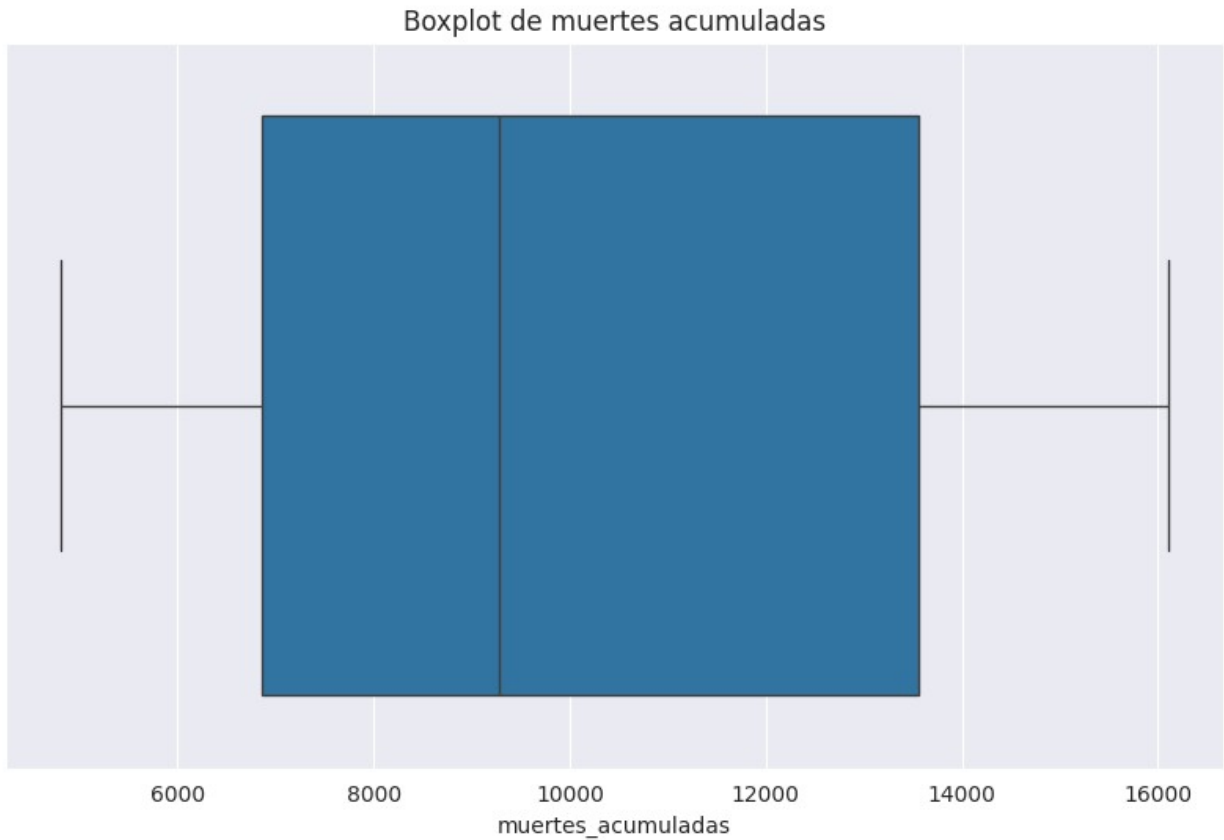
Valores unicos

```
df_nuevas_muertes.unique()

array([ 4813,  4820,  4827,  4833,  4835,  4859,  4899,  4928,  4972,
        4999,  5025,  5117,  5151,  5177,  5220,  5254,  5278,  5313,
        5343,  5389,  5420,  5456,  5465,  5469,  5486,  5513,  5543,
        5582,  5618,  5643,  5673,  5709,  5736,  5785,  5816,  5854,
        5884,  5922,  5955,  5989,  6023,  6057,  6094,  6124,  6150,
        6158,  6164,  6186,  6220,  6249,  6282,  6306,  6315,  6327,
        6334,  6348,  6374,  6393,  6402,  6412,  6427,  6435,  6467,
        6473,  6479,  6493,  6522,  6531,  6546,  6563,  6568,  6575,
        6599,  6619,  6663,  6685,  6700,  6710,  6730,  6749,  6765,
        6775,  6794,  6802,  6809,  6823,  6840,  6860,  6875,  6884,
        6891,  6894,  6903,  6928,  6955,  6977,  7001,  7016,  7020,
        7057,  7089,  7120,  7160,  7190,  7221,  7241,  7279,  7309,
        7345,  7374,  7395,  7410,  7428,  7453,  7478,  7500,  7524,
        7543,  7558,  7578,  7606,  7642,  7677,  7695,  7717,  7733,
        7736,  7776,  7815,  7832,  7845,  7872,  7887,  7893,  7913,
        7928,  7952,  7977,  7996,  8000,  8009,  8022,  8058,  8070,
        8094,  8121,  8137,  8165,  8183,  8214,  8238,  8258,  8280,
        8294,  8305,  8331,  8361,  8388,  8416,  8439,  8455,  8465,
        8500,  8527,  8549,  8595,  8645,  8694,  8735,  8785,  8845,
        8894,  8960,  9008,  9050,  9089,  9147,  9215,  9282,  9350,
        9419,  9461,  9498,  9548,  9573,  9609,  9642,  9688,  9713,
        9721,  9756,  9798,  9834,  9871,  9898,  9914,  9927, 10005,
       10016, 10029, 10063, 10093, 10100, 10112, 10176, 10224, 10248,
       10339, 10413, 10448, 10483, 10524, 10578, 10644, 10714, 10771,
       10808, 10845, 10898, 10960, 11006, 11085, 11189, 11219, 11273,
       11339, 11372, 11450, 11500, 11516, 11552, 11615, 11694, 11768,
       11818, 11858, 11886, 11926, 12007, 12098, 12155, 12203, 12265,
       12315, 12388, 12468, 12544, 12665, 12710, 12754, 12795, 12859,
       12907, 12947, 13040, 13061, 13115, 13185, 13241, 13283, 13331,
       13375, 13394, 13453, 13525, 13564, 13625, 13700, 13730, 13750,
       13794, 13851, 13902, 13938, 13986, 14027, 14066, 14118, 14177,
       14204, 14236, 14266, 14330, 14387, 14436, 14515, 14566, 14615,
       14655, 14748, 14797, 14862, 14928, 15020, 15094, 15137, 15179,
       15222, 15281, 15338, 15389, 15474, 15520, 15578, 15627, 15675,
       15714, 15730, 15749, 15775, 15802, 15837, 15840, 15849, 15855,
       15870, 15885, 15898, 15917, 15928, 15943, 15956, 15967, 15972,
       15980, 15986, 15992, 15999, 16008, 16020, 16029, 16036, 16044,
       16048, 16051, 16057, 16061, 16065, 16067, 16072, 16077, 16081,
       16085, 16089, 16095, 16096, 16097, 16098, 16099, 16102])
```

Grafica de caja

```
plt.figure(figsize=(10, 6))
ax = sns.boxplot(x=df_nuevas_muertes)
ax.set_title('Boxplot de muertes acumuladas')
plt.show()
```

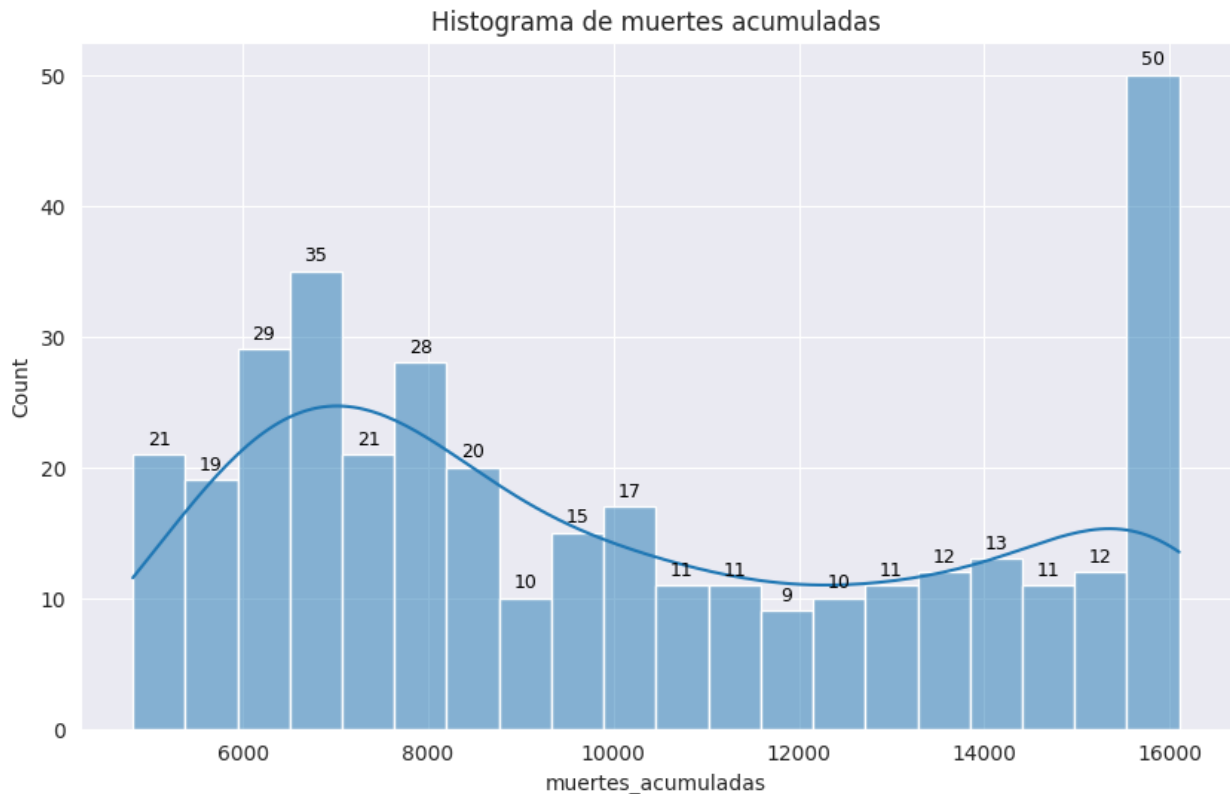


Histograma

```
# Histograma
plt.figure(figsize=(10, 6))
ax = sns.histplot(df_nuevas_muertes, kde=True, bins=20)

for p in ax.patches:
    ax.annotate(f'{p.get_height():.0f}', (p.get_x() + p.get_width() / 2.,
    p.get_height()),
    ha='center', va='baseline', fontsize=9, color='black',
    xytext=(0, 5),
    textcoords='offset points')

ax.set_title('Histograma de muertes acumuladas')
plt.show()
```



Analisis sobre datos cuantitativos sobre muertes acumuladas

En el diagrama de caja podemos observar que no hay datos atípicos. En el histograma se puede observar cuantos días pasaron sin cambios en las muertes acumuladas.

Variable: Poblacion de municipios

```
df_municipios_con_poblacion =
df_muertes_municipio.drop_duplicates(subset='codigo_municipio')
df_poblacion_municipios = df_municipios_con_poblacion['poblacion']
df_poblacion_municipios
```

0	1205668
365	86150
730	87912
1095	8317
1460	73811
...	
121180	29121
121545	50348
121910	11057
122275	13786

```
122640      23311
Name: poblacion, Length: 337, dtype: int64
```

Describe

```
df_poblacion_municipios.describe()
```

```
count      3.370000e+02
mean       4.997144e+04
std        8.158246e+04
min        2.563000e+03
25%        1.746500e+04
50%        3.091200e+04
75%        5.852600e+04
max        1.205668e+06
Name: poblacion, dtype: float64
```

Valores unicos

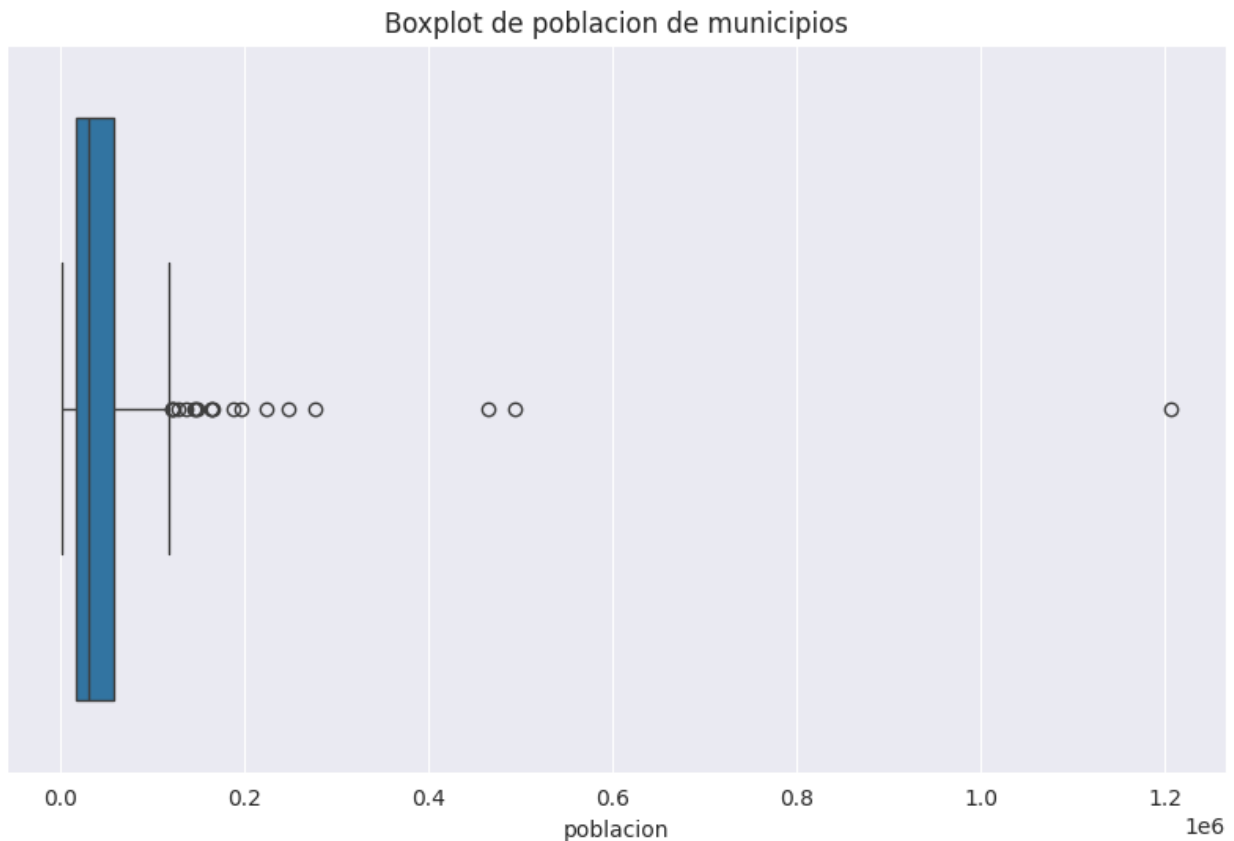
```
df_poblacion_municipios.unique()
```

```
array([1205668,  86150,  87912,  8317,  73811, 121721,  65594,
        494561,  54623, 276836, 38386, 17957,  61664, 147604,
        464528, 165026, 145417, 27001, 12577,  54588,   7851,
        13508,  13535,  41997, 22012,  58338,  23160,  19389,
        45249,  14121,  36409,   8519,  28445,  17116,  12674,
        26472,  40083,  15231, 27787,  12438,   4480, 112778,
        32963,  92103,  58174, 23166, 108893,  69560,  10996,
        39337,  10646,  28043, 41304,  34591,  19693,  28473,
        22618, 166078, 123979, 25479,  25677,  47441,  62895,
        51024,  17234,  66314, 19778,  70118,  17923,  80234,
        17465,  46489,  60376, 21088,  25859,  13620,  26686,
        12599,  63897,  36117, 26560,  21906,  17569,  29285,
        38981,  98886,   5128,   2563,  21695,  91523,  60200,
        10200,  14640,  15840,  15639,  31581,   7817,   7950,
        2911,  13123,  11600, 44025, 117483,  42354,  64224,
        29356, 136909,  58526, 25461,  12374, 196867,  20851,
        38219,  34866,   8766, 27307,  18896,   9607,  60791,
        8658,  20245,  32771,  16679,  45229,  14522,  15414,
        55290,   9361,  30973, 117579,  45162,  23851,  24199,
        20299,  83448,  37283,  24289,  16817,  11870,  48188,
        13811,  26350,  22048,  64441,  12261,   7945,  67994,
        41984,  27567,   9238,  12569,   8724,  11980,  30776,
        25716, 108405,  30444,  15835,  12905,  26191,  55133,
        38006,  41552,  42601,  48644,  84350,  20579,  79867,
        51975,  82876,  81005,  16786,  62329,  43512,  18128,
        31500,  51299,  19080, 108134,  32104,  45434,  12573,
        51253,  25475,  23394,  23021,  24580,  22423,  17811,
        26714,  13785,  13803,  33764, 129232,  91954,  21855,
        65178,  49299,  40493,  42375,  61547,  54382,  42667,
```


46629,	59788,	35071,	18222,	38303,	25165,	50171,
57277,	43622,	41035,	11635,	20384,	30205,	18291,
25332,	122466,	57977,	12330,	10341,	13971,	10225,
17393,	6861,	97974,	32038,	13294,	39269,	50814,
148943,	7462,	30134,	37419,	46350,	35616,	104861,
75979,	27522,	72955,	54845,	15586,	42342,	107265,
9652,	70066,	35895,	45937,	59496,	14380,	10229,
27591,	64292,	224109,	34981,	79308,	41224,	23858,
50093,	79777,	97296,	248880,	60355,	26146,	69107,
104428,	28890,	76047,	44819,	39350,	43067,	7103,
52439,	34823,	83375,	16339,	23499,	29659,	76128,
97003,	30017,	66883,	37915,	15225,	109753,	79360,
81385,	106156,	63841,	65730,	10859,	23030,	49410,
20032,	13071,	14440,	7544,	38270,	11948,	13384,
117014,	9088,	18085,	73093,	60848,	29283,	55723,
12509,	29377,	13124,	23435,	188522,	76020,	26170,
8139,	21416,	28097,	45870,	163665,	25189,	30812,
17322,	51770,	17833,	19514,	6706,	8272,	10340,
36047,	30912,	29121,	50348,	11057,	13786,	23311])

Grafica de caja

```
plt.figure(figsize=(10, 6))
ax = sns.boxplot(x=df_poblacion_municipios)
ax.set_title('Boxplot de poblacion de municipios')
plt.show()
```

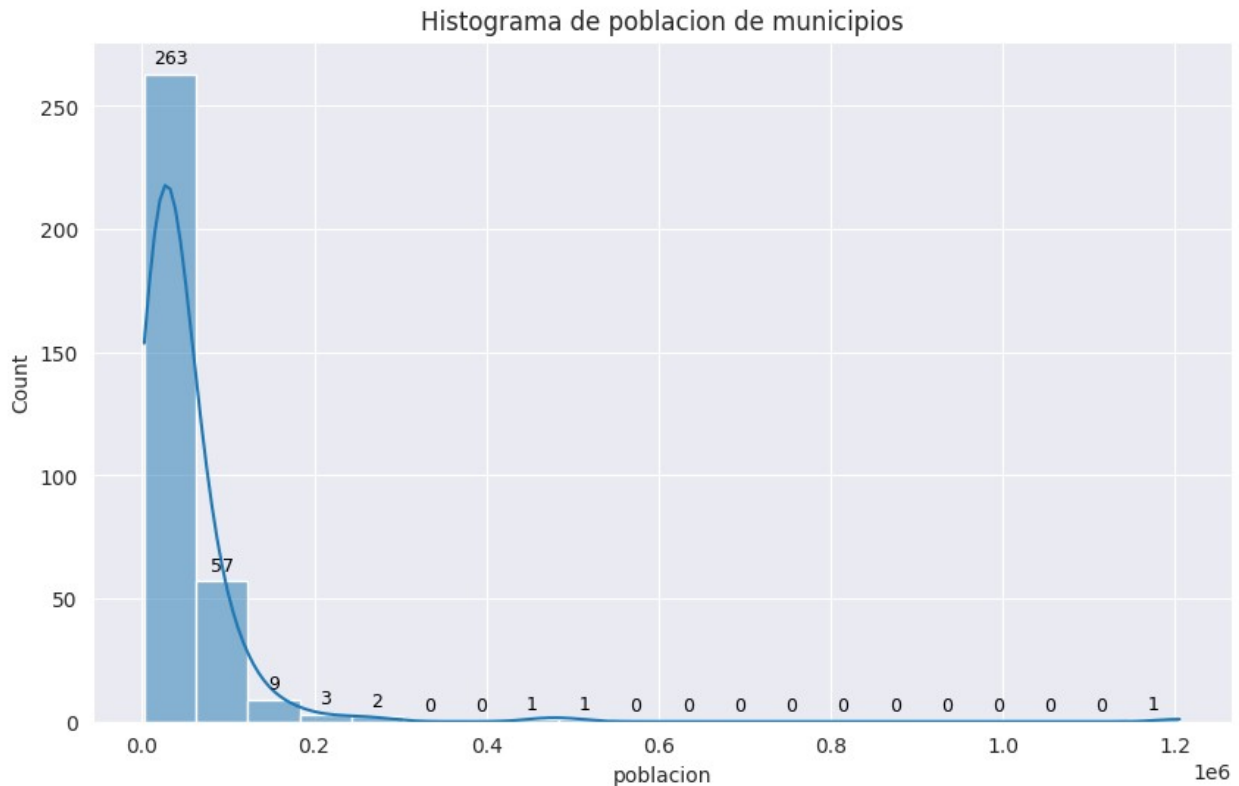


Histograma

```
# Histograma
plt.figure(figsize=(10, 6))
ax = sns.histplot(df_poblacion_municipios, kde=True, bins=20)

for p in ax.patches:
    ax.annotate(f'{p.get_height():.2f}', (p.get_x() + p.get_width() / 2.,
    p.get_height()),
    ha='center', va='baseline', fontsize=9, color='black',
    xytext=(0, 5),
    textcoords='offset points')

ax.set_title('Histograma de poblacion de municipios')
plt.show()
```



Removiendo outliers

```
Q1 = df_poblacion_municipios.quantile(0.25)
Q3 = df_poblacion_municipios.quantile(0.75)
IQR = Q3 - Q1

poblacion_sin_outliers = (df_poblacion_municipios < Q1 - 1.5 * IQR) |
(df_poblacion_municipios > Q3 + 1.5 * IQR)
poblacion_sin_outliers =
df_poblacion_municipios[~poblacion_sin_outliers]
poblacion_sin_outliers.describe()
```

count	318.000000
mean	37639.512579
std	26515.215046
min	2563.000000
25%	16891.750000
50%	29320.500000
75%	51652.250000
max	117579.000000
Name: poblacion, dtype: float64	

Nuevo histograma sin outliers

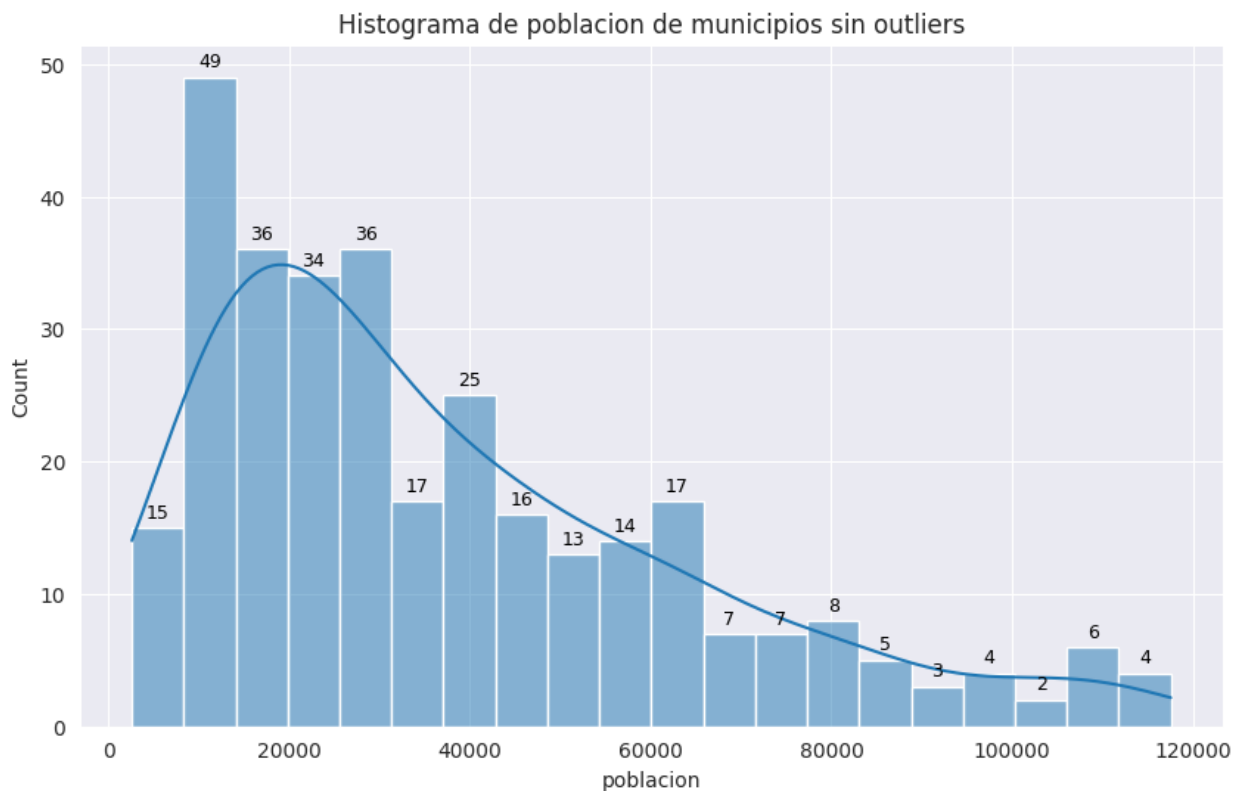
```
plt.figure(figsize=(10, 6))
ax = sns.histplot(poblacion_sin_outliers, kde=True, bins=20)
```

```

for p in ax.patches:
    ax.annotate(f'{p.get_height()}', (p.get_x() + p.get_width() / 2.,
    p.get_height()),
                ha='center', va='baseline', fontsize=9, color='black',
    xytext=(0, 5),
                textcoords='offset points')

ax.set_title('Histograma de poblacion de municipios sin outliers')
plt.show()

```



Analisis sobre datos cuantitativos sobre poblacion de municipios

En el diagrama de caja podemos observar que hay datos atipos muy altos, lo que provoca que el histograma no sea muy claro. Una vez removidos estos datos atipicos, podemos observar que la mayoria de los municipios tienen una poblacion entre 0 y 65000.

Datos Cualitativos

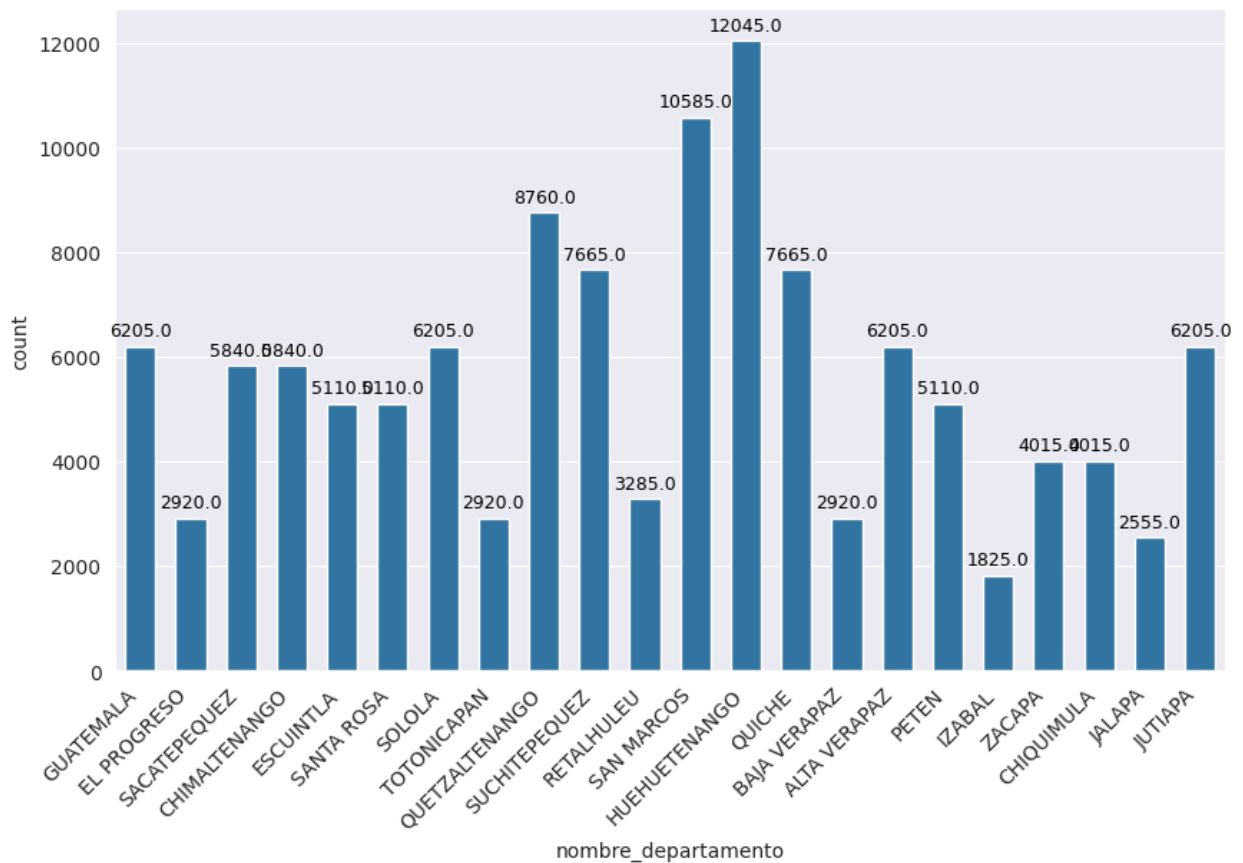
Variable: Departamentos

Diagrama de barras

```
plt.figure(figsize=(10, 6))
ax = sns.countplot(x='nombre_departamento', data=df_muertes_municipio,
width=0.6)

for p in ax.patches:
    ax.annotate(f'{p.get_height()}', (p.get_x() + p.get_width() / 2.,
p.get_height()),
                ha='center', va='baseline', fontsize=9, color='black',
xytext=(0, 5),
                textcoords='offset points')

ax.set_xticklabels(ax.get_xticklabels(), rotation=45, ha="right")
plt.show()
```



Analisis sobre datos cualitativos sobre departamentos

Se puede observar que los dos departamentos con mas registros son los departamentos de Huehuetenango y San Marcos, esto se debe a que estos dos departamentos son los que tienen mas municipios en el pais. Se debe tomar en cuenta que esta grafica no representa la cantidad de muertes por departamento, sino la cantidad de registros por departamento.

Variable: Municipios

Diagrama de barras Guatemala

```
df_depto_guatemala =
df_muertes_municipio[df_muertes_municipio['nombre_departamento'] ==
'GUATEMALA']
plt.figure(figsize=(10, 6))
ax = sns.countplot(x='nombre_municipio', data=df_depto_guatemala,
width=0.6)

for p in ax.patches:
    ax.annotate(f'{p.get_height()}', (p.get_x() + p.get_width() / 2.,
p.get_height()),
                ha='center', va='baseline', fontsize=9, color='black',
xytext=(0, 5),
                textcoords='offset points')

ax.set_xticklabels(ax.get_xticklabels(), rotation=45, ha="right")
plt.show()
```

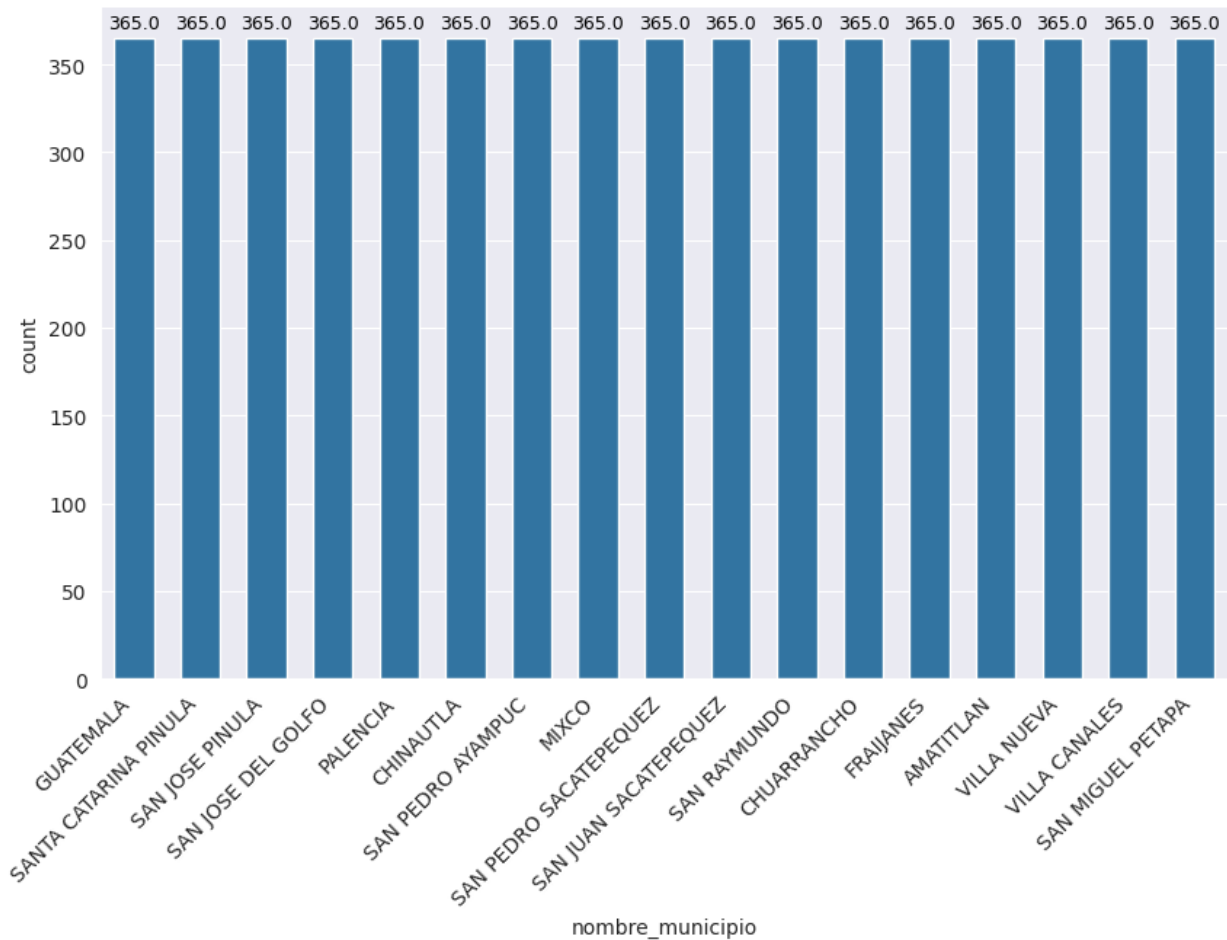


Diagrama de barras Quetzaltenango

```
plt.figure(figsize=(10, 8))
df_depto_xela =
df_muertes_municipio[df_muertes_municipio['nombre_departamento'] ==
'QUETZALTENANGO']
ax = sns.countplot(x='nombre_municipio', data=df_depto_xela,
width=0.6)

for p in ax.patches:
    ax.annotate(f'{p.get_height()}', (p.get_x() + p.get_width() / 2.,
p.get_height()),
                ha='center', va='baseline', fontsize=8, color='black',
xytext=(0, 5),
                textcoords='offset points')

ax.set_xticklabels(ax.get_xticklabels(), rotation=45, ha="right")
plt.show()
```

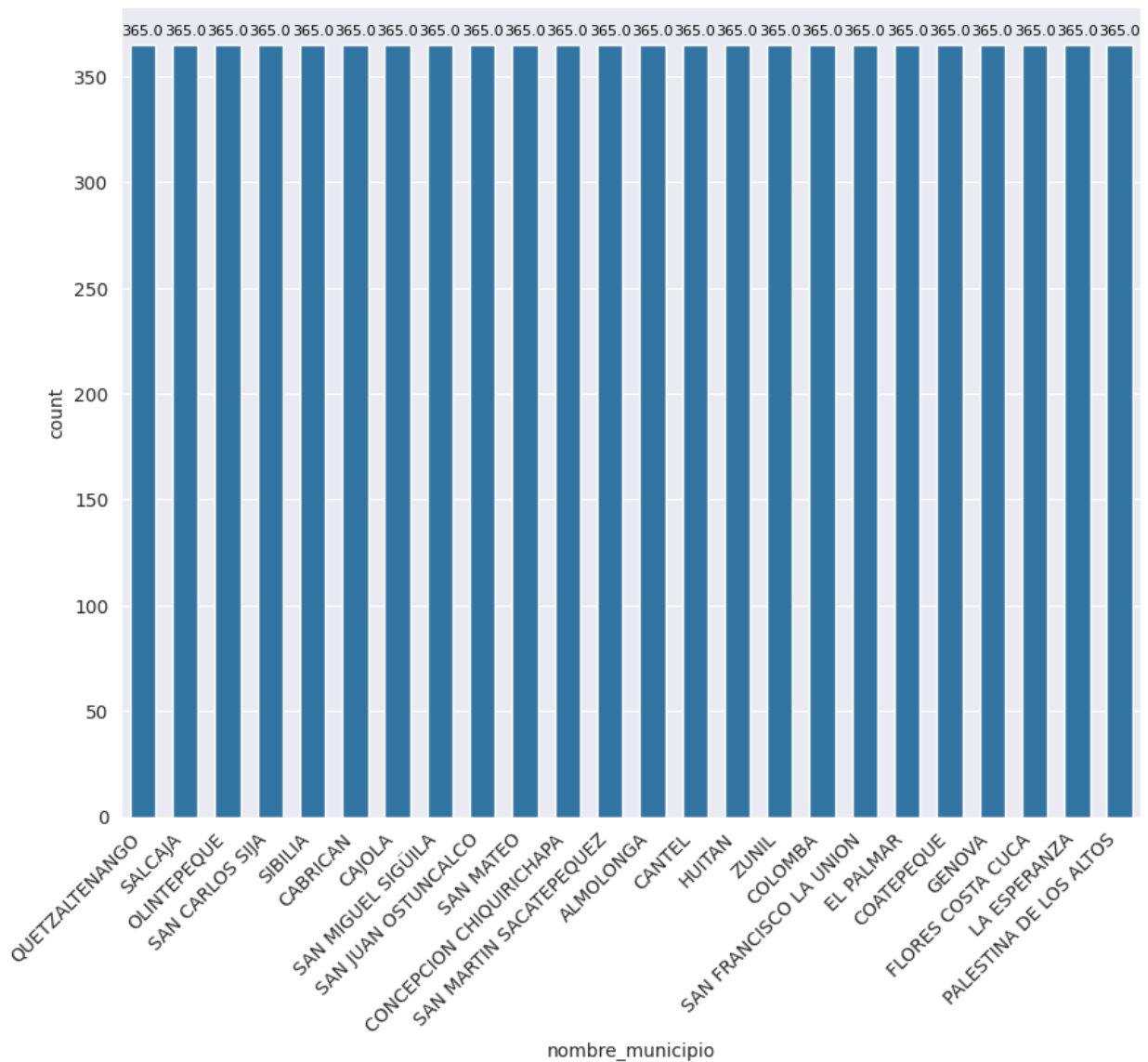


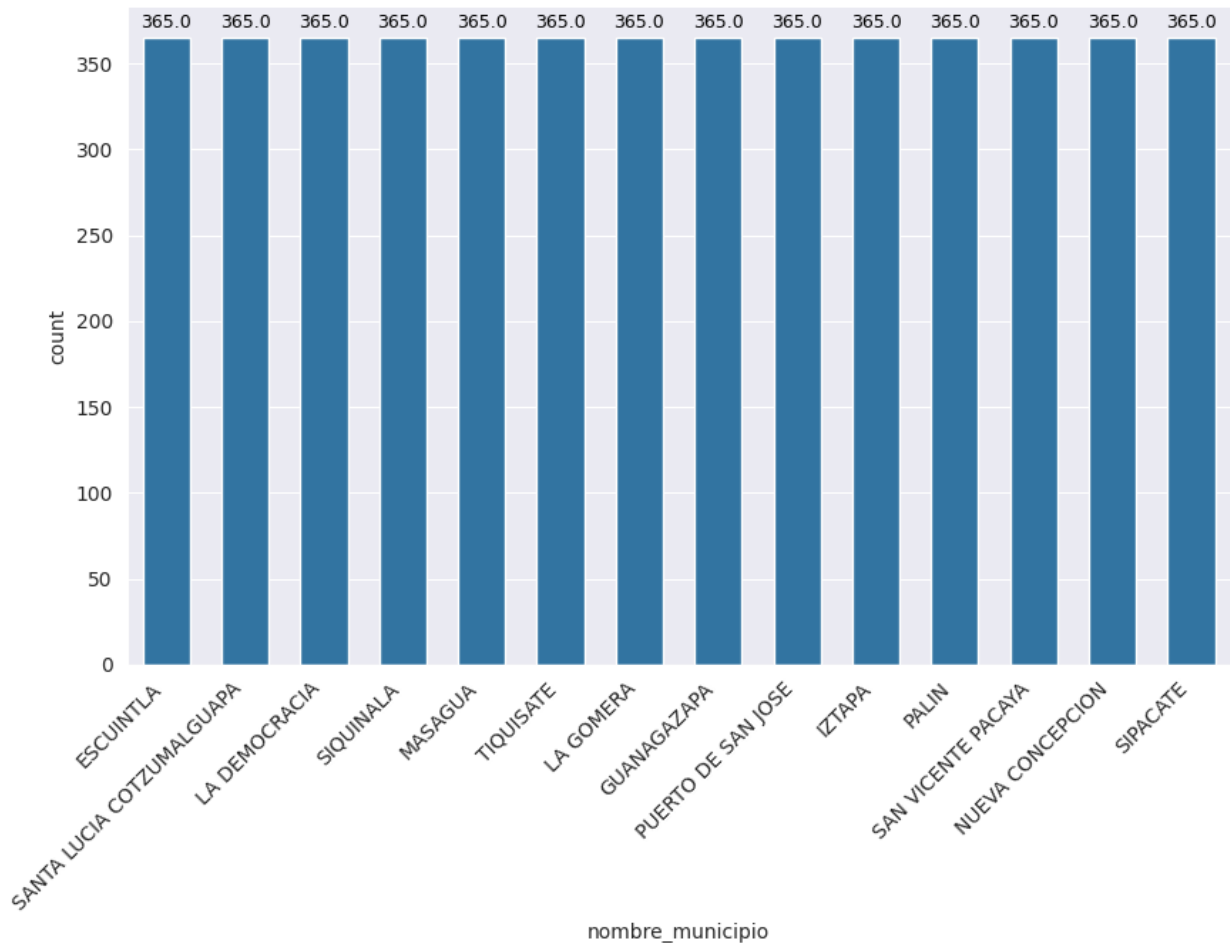
Diagrama de barras Escuintla

```
plt.figure(figsize=(10, 6))
df_depto_escuintla =
df_muertes_municipio[df_muertes_municipio['nombre_departamento'] ==
'ESCUINTLA']
ax = sns.countplot(x='nombre_municipio', data=df_depto_escuintla,
width=0.6)

for p in ax.patches:
    ax.annotate(f'{p.get_height()}', (p.get_x() + p.get_width() / 2.,
p.get_height()),
                ha='center', va='baseline', fontsize=9, color='black',
xytext=(0, 5),
                textcoords='offset points')
```



```
ax.set_xticklabels(ax.get_xticklabels(), rotation=45, ha="right")
plt.show()
```



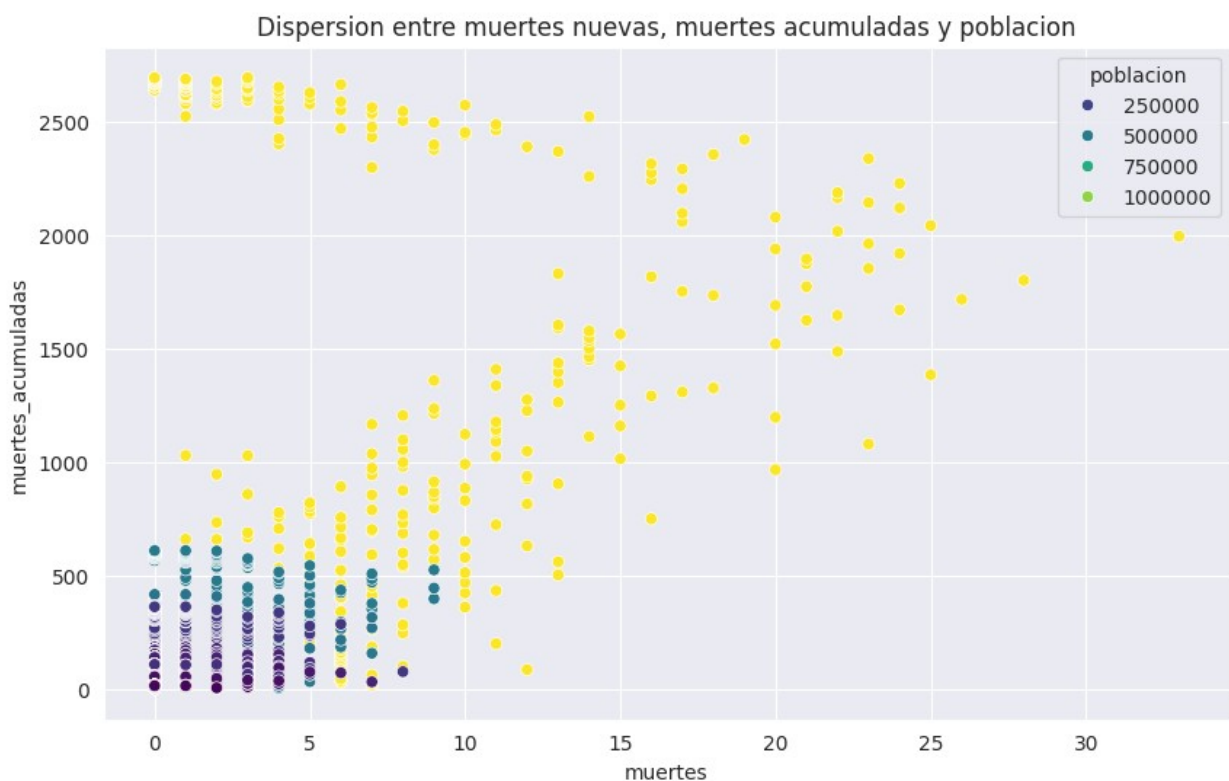
Analisis sobre datos cualitativos sobre municipios

El conteo de los registros por municipio tienen 365 para todos ya que los datos proporcionados fueron datos de muertes por fecha para todo un año en específico, por lo que solo se realizó la grafica de los municipios para los siguientes departamentos: Guatemala Quetzaltengo Escuintla

Analisis EDA MultiVariable

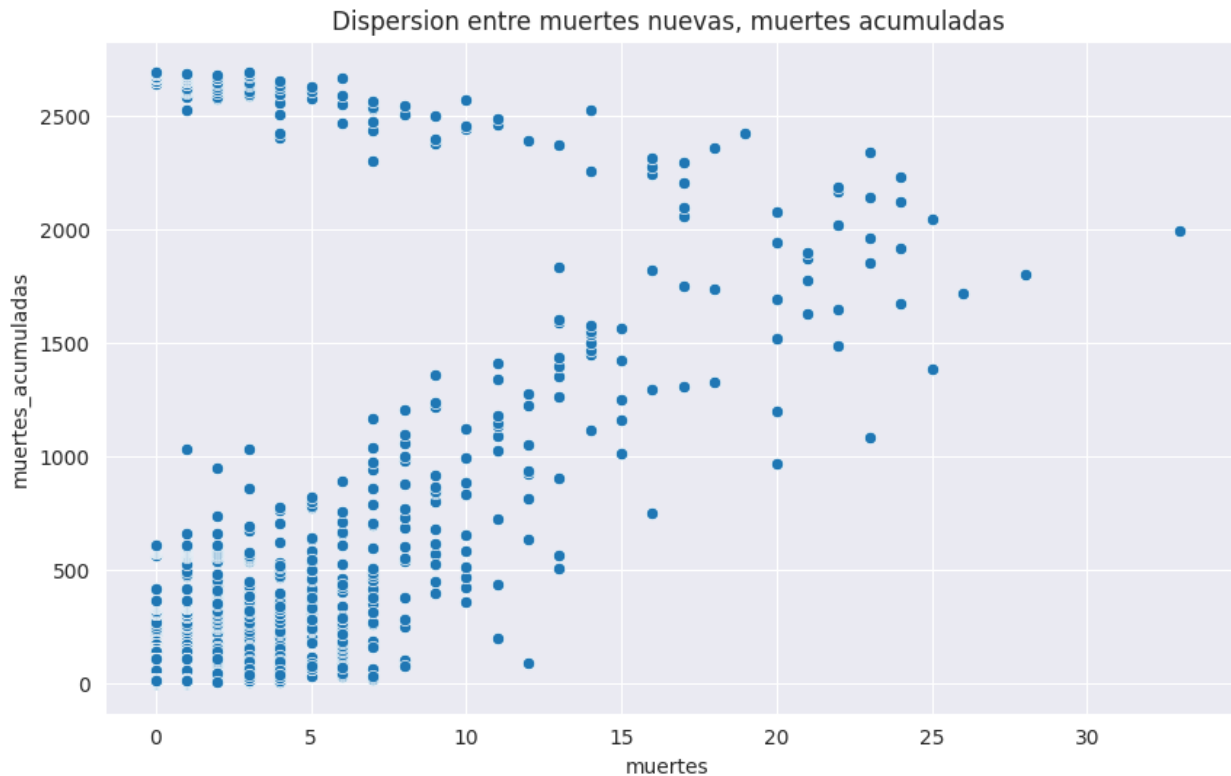
Dispersion entre variables muertes, muertes acumuladas y poblacion

```
plt.figure(figsize=(10, 6))
ax = sns.scatterplot(x='muertes', y='muertes_acumuladas',
data=df_muertes_municipio, hue='poblacion', palette='viridis')
ax.set_title('Dispersion entre muertes nuevas, muertes acumuladas y
poblacion')
plt.show()
```



Dispersion entre variables muertes, muertes acumuladas

```
plt.figure(figsize=(10, 6))
ax = sns.scatterplot(x='muertes', y='muertes_acumuladas',
data=df_muertes_municipio, palette='viridis')
ax.set_title('Dispersion entre muertes nuevas, muertes acumuladas')
plt.show()
```



Analisis sobre dispersion entre variables muertes, muertes acumuladas y poblacion

En la grafica se puede observar que cantidades menores de muertes se dieron mas seguido al inicio de la pandemia es decir cuando las muertes acumuladas eran menores, mientras que las cantidades mas grandes de muertes se dieron cuando las muertes acumuladas eran mayores.

Mapa de calor entre variables muertes, muertes acumuladas y poblacion

```
plt.figure(figsize=(10, 6))
ax = sns.heatmap(df_muertes_municipio[['muertes',
'muertes_acumuladas', 'poblacion']].corr(), annot=True)
ax.set_title('Mapa de calor entre muertes nuevas, muertes acumuladas y
poblacion')
plt.show()
```

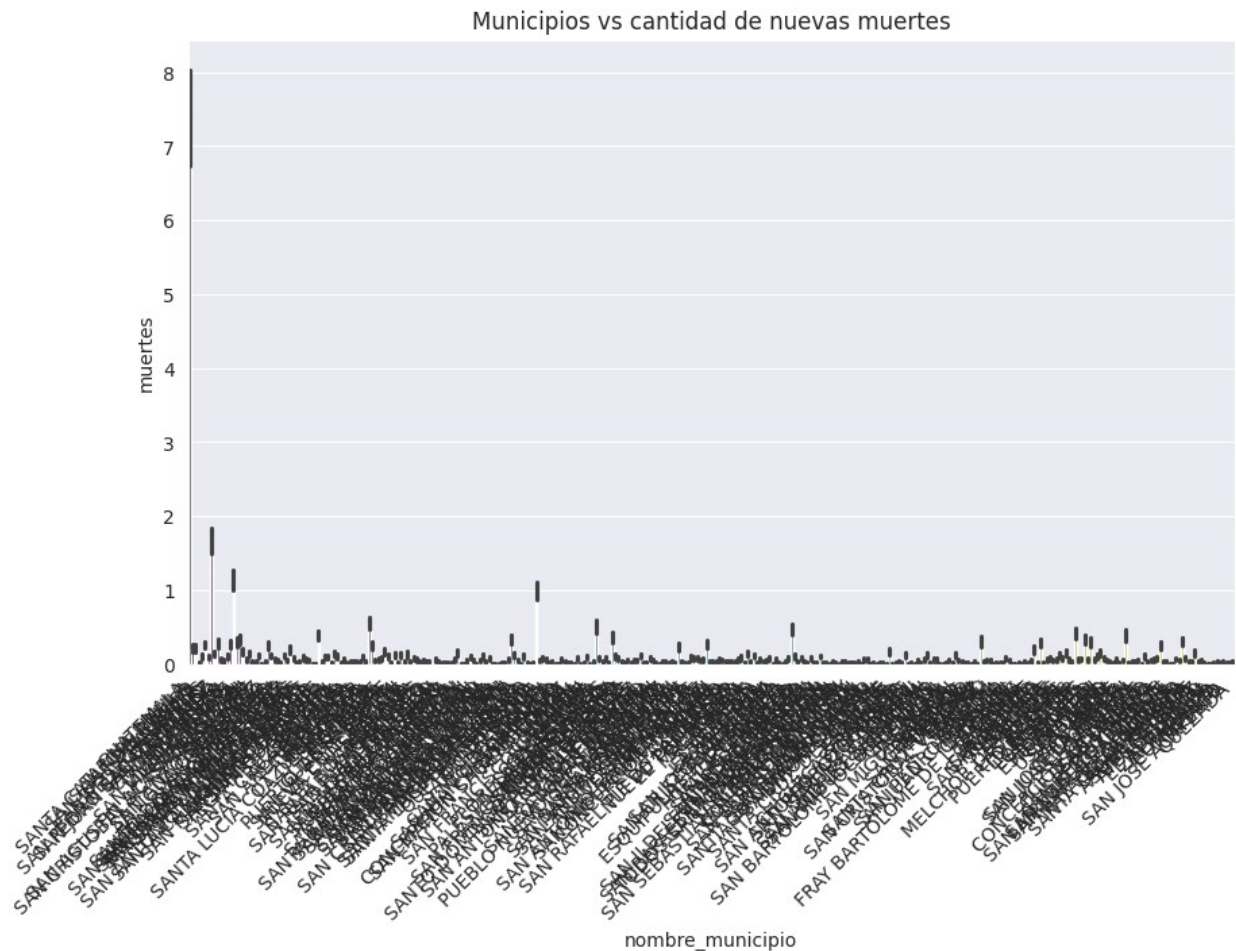


Analisis sobre mapa de calor entre variables muertes, muertes acumuladas y poblacion

En el mapa podemos observar que la poblacion tiene una correlacion alta con ambas variables, sin embargo la correlacion mayor es con las muertes acumuladas. La correlacion entre muertes y muertes acumuladas es moderada. Estas correlaciones nos indican que a mayor poblacion mayor cantidad de muertes acumuladas y nuevas muertes.

Gráfico de barras: Municipios vs cantidad de nuevas muertes

```
plt.figure(figsize=(10, 6))
ax = sns.barplot(x='nombre_municipio', y='muertes',
data=df_muertes_municipio, palette='viridis')
ax.set_xticklabels(ax.get_xticklabels(), rotation=45, ha="right")
ax.set_title('Municipios vs cantidad de nuevas muertes')
plt.show()
```

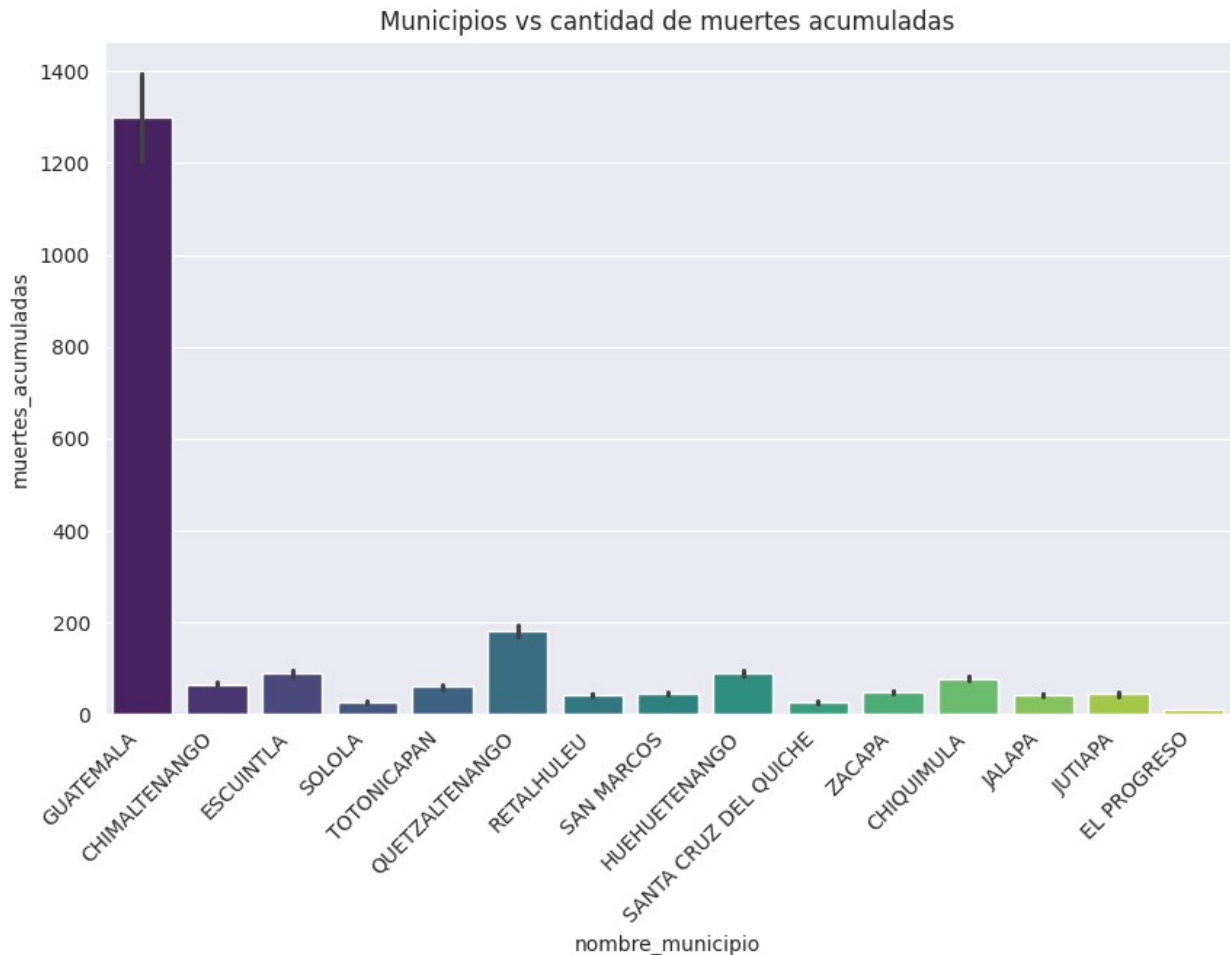


Grafica de barras entre municipios (cabeceras) y cantidad de muertes acumuladas

Codigos de municipios que son cabeceras:

101,401,501,701,801,901,1101,1201,1301,1401,1901,2001,2101,2201,2202

```
df_municipios_cabeceras =
df_muertes_municipio[df_muertes_municipio['codigo_municipio'].isin([10
1,401,501,701,801,901,1101,1201,1301,1401,1901,2001,2101,2201,2202])]
plt.figure(figsize=(10, 6))
ax = sns.barplot(x='nombre_municipio', y='muertes_acumuladas',
data=df_municipios_cabeceras, palette='viridis')
ax.set_xticklabels(ax.get_xticklabels(), rotation=45, ha="right")
ax.set_title('Municipios vs cantidad de muertes acumuladas')
plt.show()
```



Analisis sobre grafica de barras entre municipios y cantidad de muertes acumuladas

Se puede observar que de los municipios que son cabeceras el que mayor cantidad de muertes acumuladas tiene es el municipio de Guatemala.

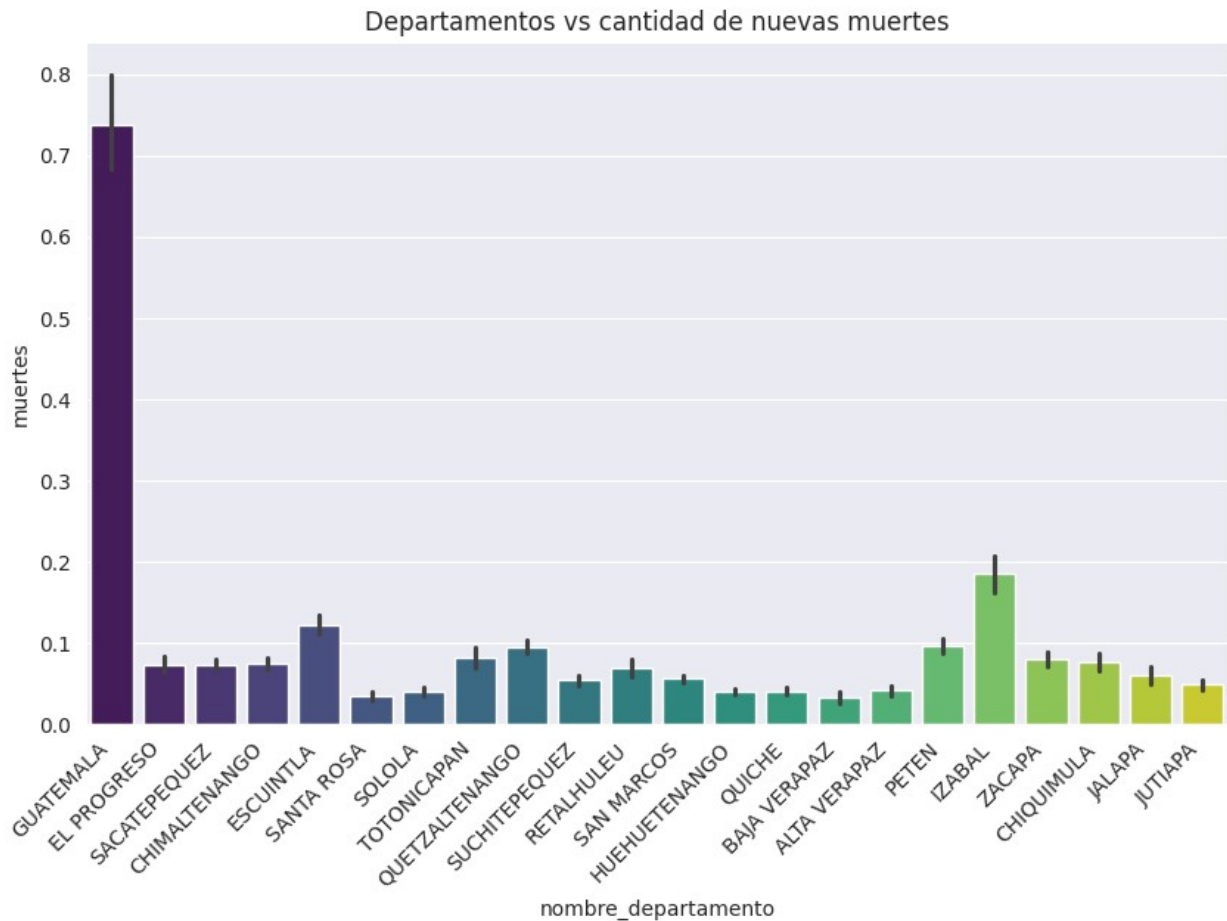
Mapa de calor entre municipios y cantidad de nuevas muertes

```
plt.figure(figsize=(10, 6))
ax = sns.heatmap(df_muertes_municipio[['muertes',
'codigo_municipio']].corr(), annot=True)
ax.set_title('Mapa de calor entre municipios y cantidad de nuevas
muertes')
plt.show()
```



Grafica de barras: Departamentos vs cantidad de nuevas muertes

```
plt.figure(figsize=(10, 6))
ax = sns.barplot(x='nombre_departamento', y='muertes',
data=df_muertes_municipio, palette='viridis')
ax.set_xticklabels(ax.get_xticklabels(), rotation=45, ha="right")
ax.set_title('Departamentos vs cantidad de nuevas muertes')
plt.show()
```



Analisis sobre grafica de barras entre departamentos y cantidad de nuevas muertes

Se puede observar que el departamento con mayor cantidad de muertes es el departamento de Guatemala, esto se debe a que es un municipio con mucha poblacion.

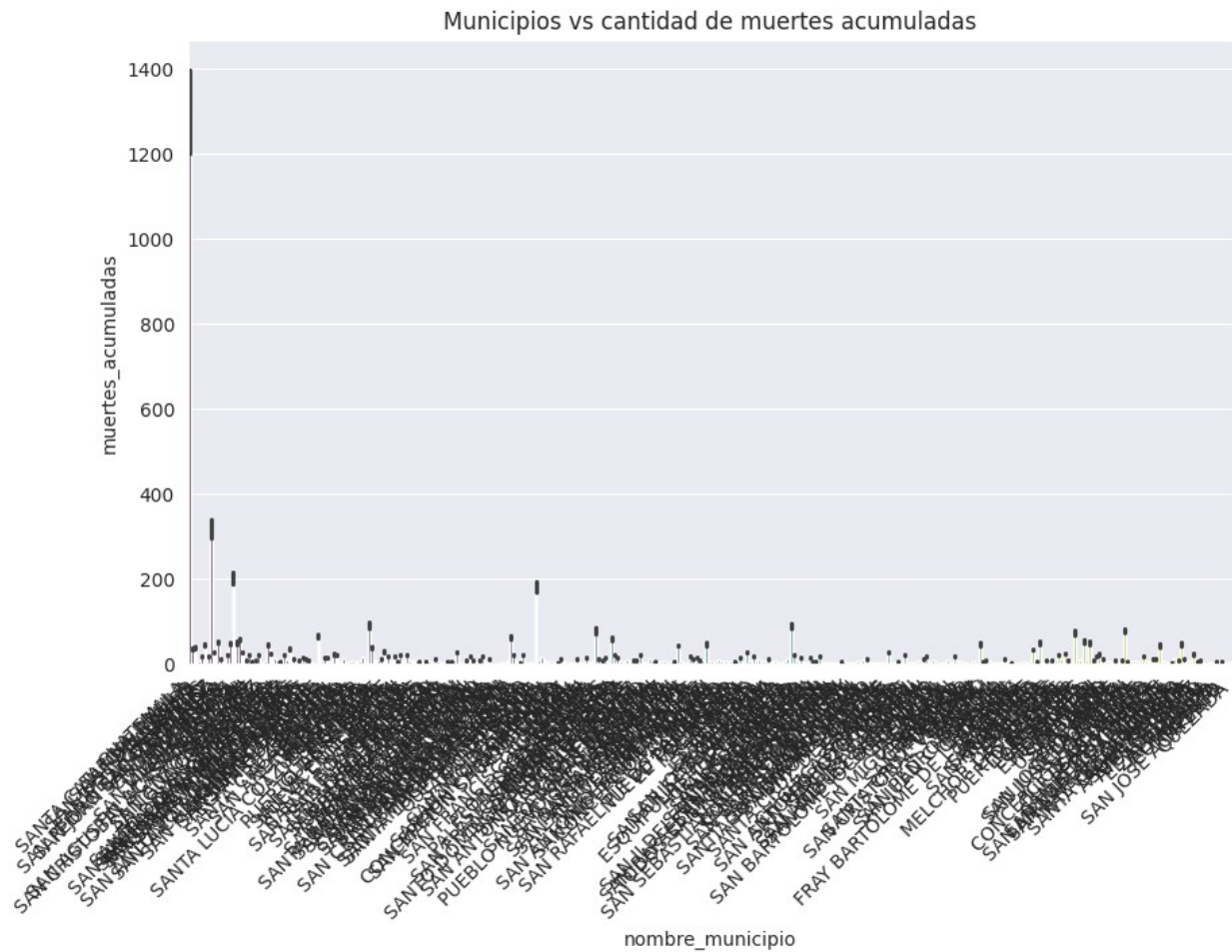
Mapa de calor entre departamentos y cantidad de nuevas muertes

```
plt.figure(figsize=(10, 6))
ax = sns.heatmap(df_muertes_municipio[['muertes',
'codigo_departamento']].corr(), annot=True)
ax.set_title('Mapa de calor entre departamentos y cantidad de nuevas
muertes')
plt.show()
```



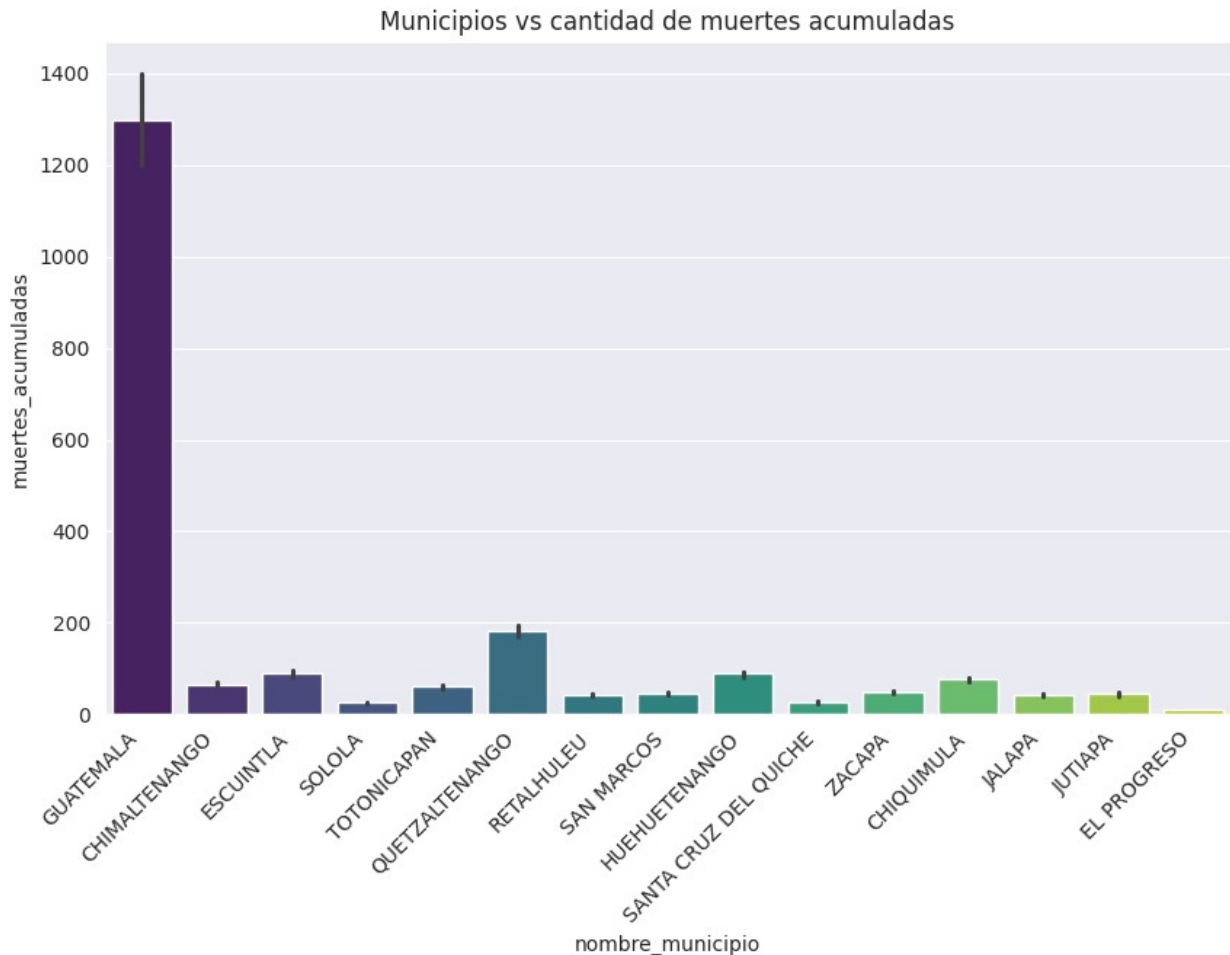

Grafica de barras municipios vs cantidad de muertes acumuladas

```
plt.figure(figsize=(10, 6))
ax = sns.barplot(x='nombre_municipio', y='muertes_acumuladas',
data=df_muertes_municipio, palette='viridis')
ax.set_xticklabels(ax.get_xticklabels(), rotation=45, ha="right")
ax.set_title('Municipios vs cantidad de muertes acumuladas')
plt.show()
```



Grafica de barras entre municipios (cabeceras) y cantidad de muertes acumuladas

```
plt.figure(figsize=(10, 6))
ax = sns.barplot(x='nombre_municipio', y='muertes_acumuladas',
data=df_municipios_cabeceras, palette='viridis')
ax.set_xticklabels(ax.get_xticklabels(), rotation=45, ha="right")
ax.set_title('Municipios vs cantidad de muertes acumuladas')
plt.show()
```



Analisis sobre grafica de barras entre municipios y cantidad de muertes acumuladas

Se puede observar que el municipio de Guatemala es el que tiene mayor cantidad de muertes_acumuladas, esto se debe a que es un municipio con mucha poblacion.

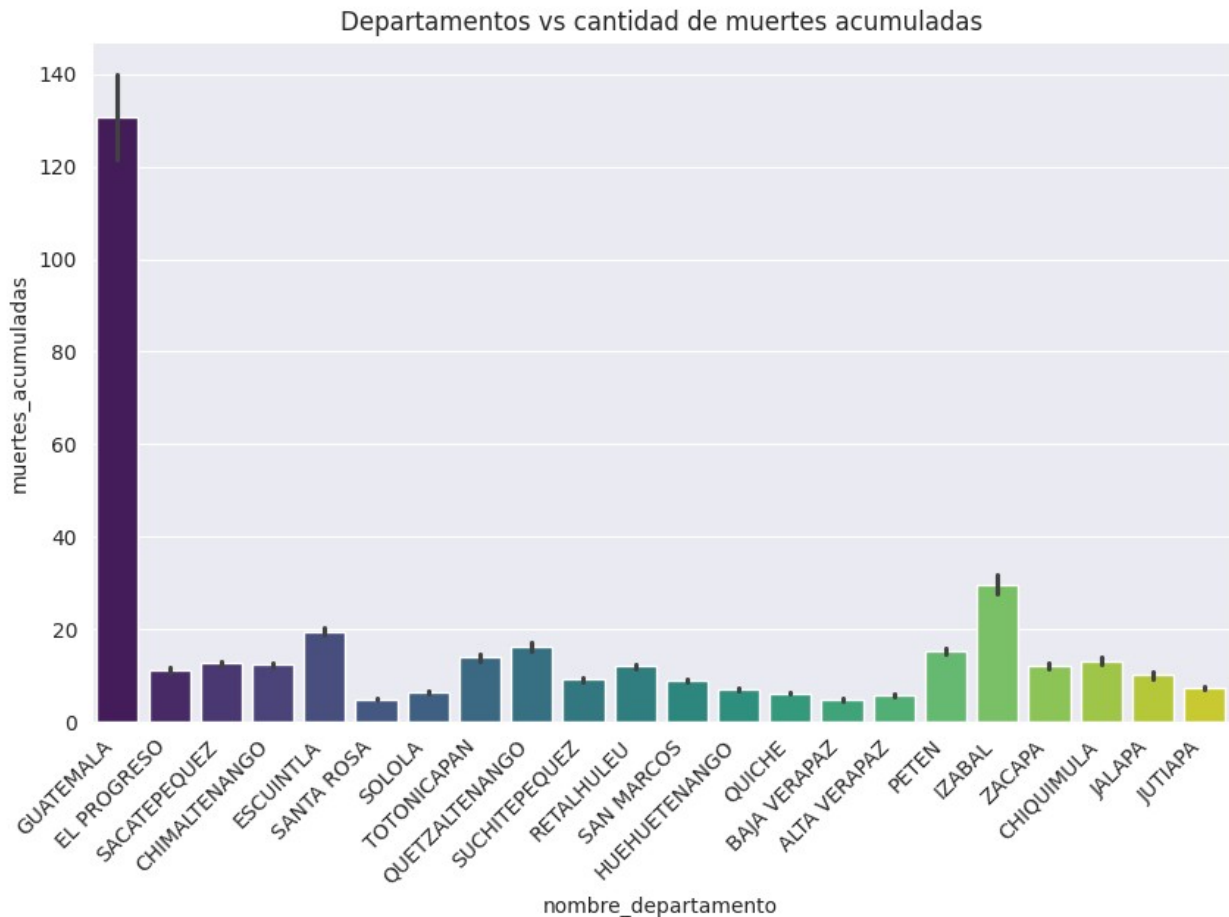
Mapa de calor entre municipios y cantidad de muertes acumuladas

```
plt.figure(figsize=(10, 6))
ax = sns.heatmap(df_muertes_municipio[['muertes_acumuladas',
'codigo_municipio']].corr(), annot=True)
ax.set_title('Mapa de calor entre municipios y cantidad de muertes
acumuladas')
plt.show()
```



Grafica de barras departamentos vs cantidad de muertes acumuladas

```
plt.figure(figsize=(10, 6))
ax = sns.barplot(x='nombre_departamento', y='muertes_acumuladas',
data=df_muertes_municipio, palette='viridis')
ax.set_xticklabels(ax.get_xticklabels(), rotation=45, ha="right")
ax.set_title('Departamentos vs cantidad de muertes acumuladas')
plt.show()
```



Analisis sobre grafica de barras entre departamentos y cantidad de muertes acumuladas

Siempre el departamento de Guatemala es el que tiene mayor cantidad de muertes acumuladas, respecto a las demas cabeceras.

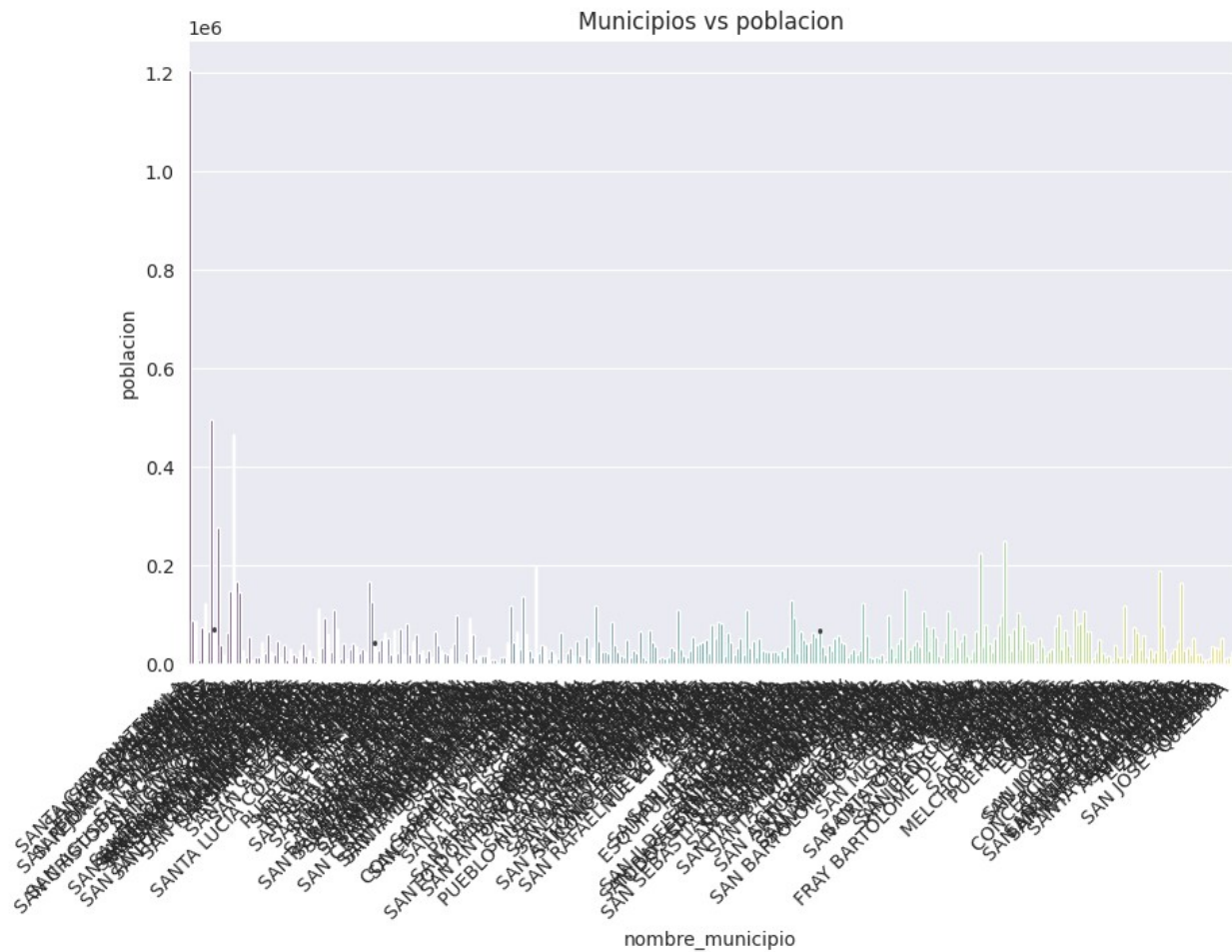
Mapa de calor entre departamentos y cantidad de muertes acumuladas

```
plt.figure(figsize=(10, 6))
ax = sns.heatmap(df_muertes_municipio[['muertes_acumuladas',
'codigo_departamento']].corr(), annot=True)
ax.set_title('Mapa de calor entre departamentos y cantidad de muertes acumuladas')
plt.show()
```



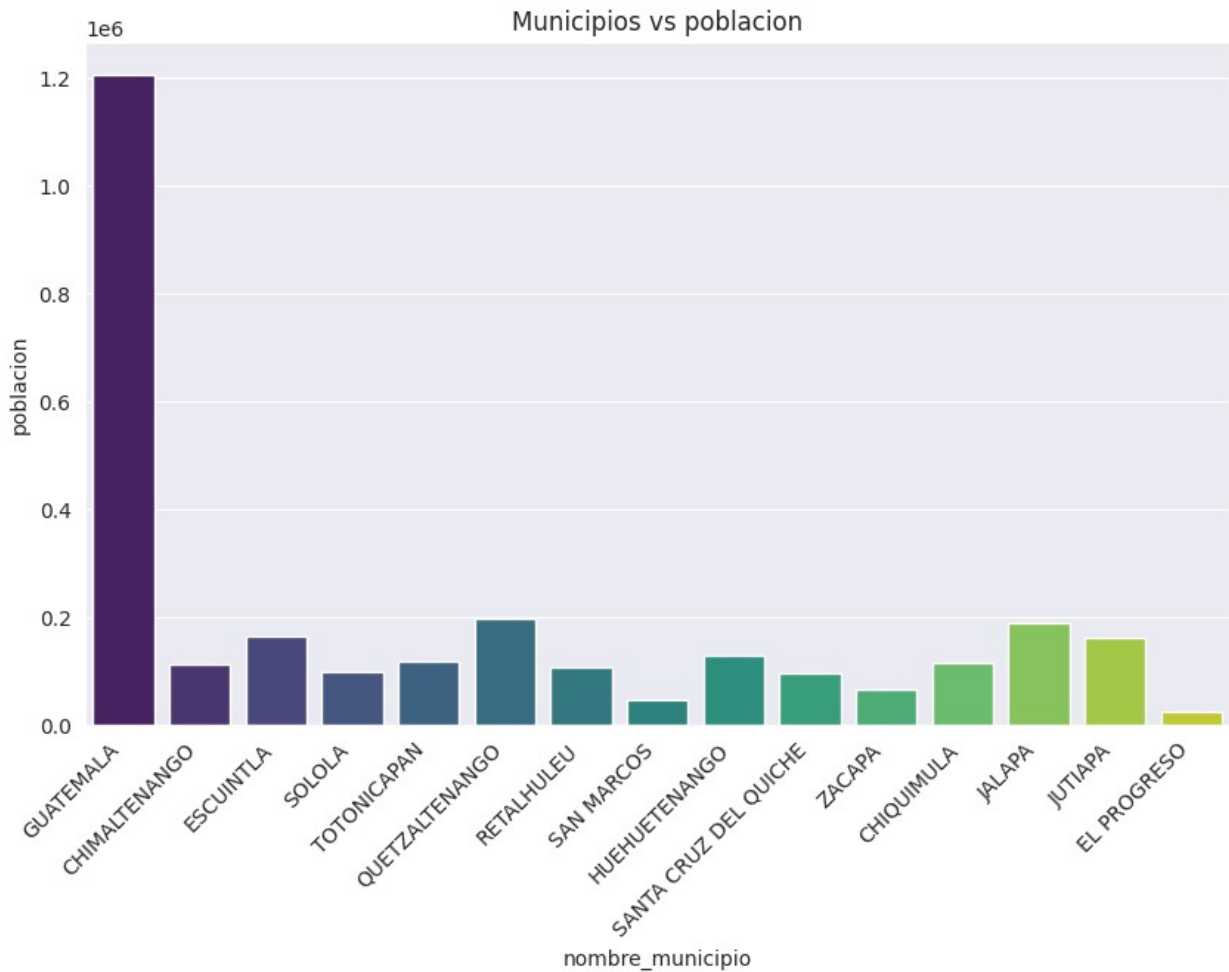
Grafica de barras municipios vs poblacion

```
plt.figure(figsize=(10, 6))
ax = sns.barplot(x='nombre_municipio', y='poblacion',
data=df_muertes_municipio, palette='viridis')
ax.set_xticklabels(ax.get_xticklabels(), rotation=45, ha="right")
ax.set_title('Municipios vs poblacion')
plt.show()
```



Grafica de barras entre municipios (cabeceras) y poblacion

```
plt.figure(figsize=(10, 6))
ax = sns.barplot(x='nombre_municipio', y='poblacion',
data=df_municipios_cabeceras, palette='viridis')
ax.set_xticklabels(ax.get_xticklabels(), rotation=45, ha="right")
ax.set_title('Municipios vs poblacion')
plt.show()
```

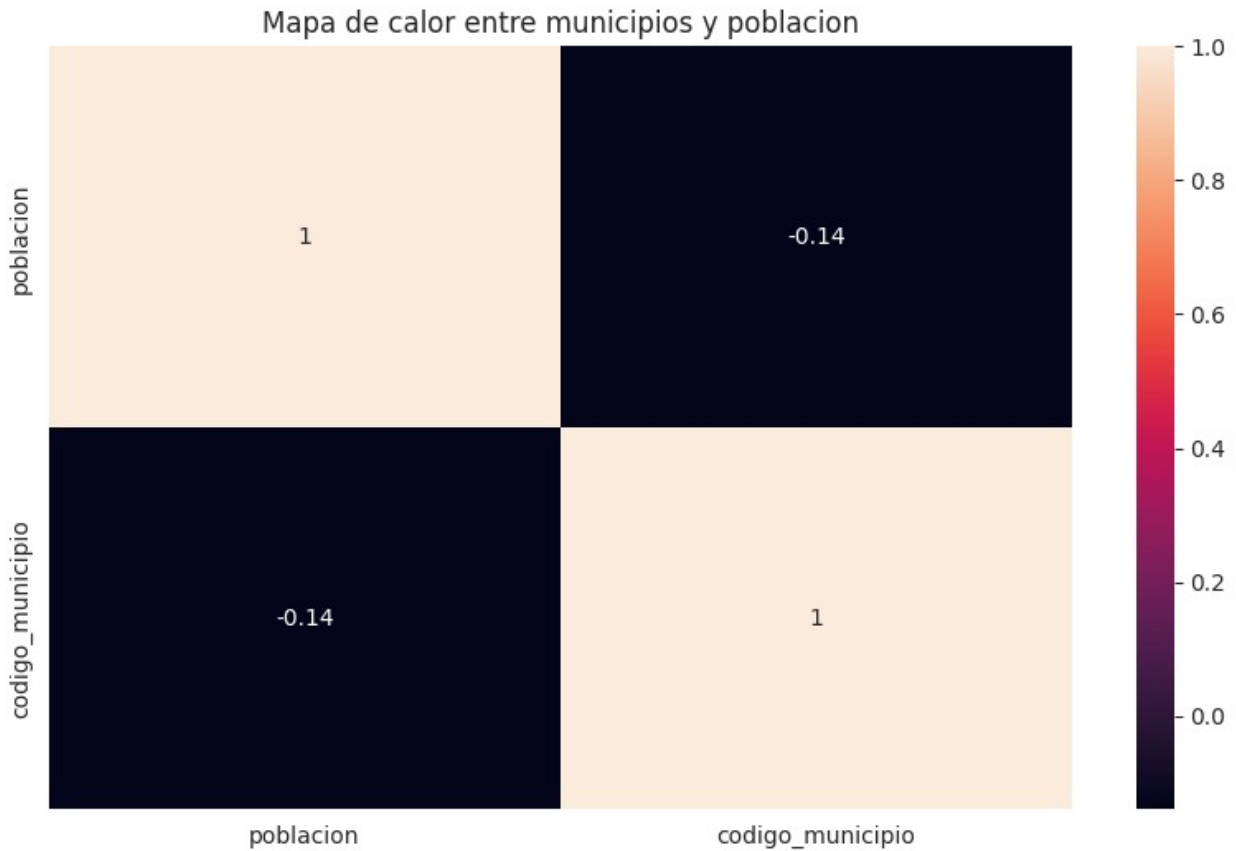


Analisis sobre grafica de barras entre municipios y poblacion

Se puede observar que el municipio de Guatemala es el que tiene mayor poblacion.

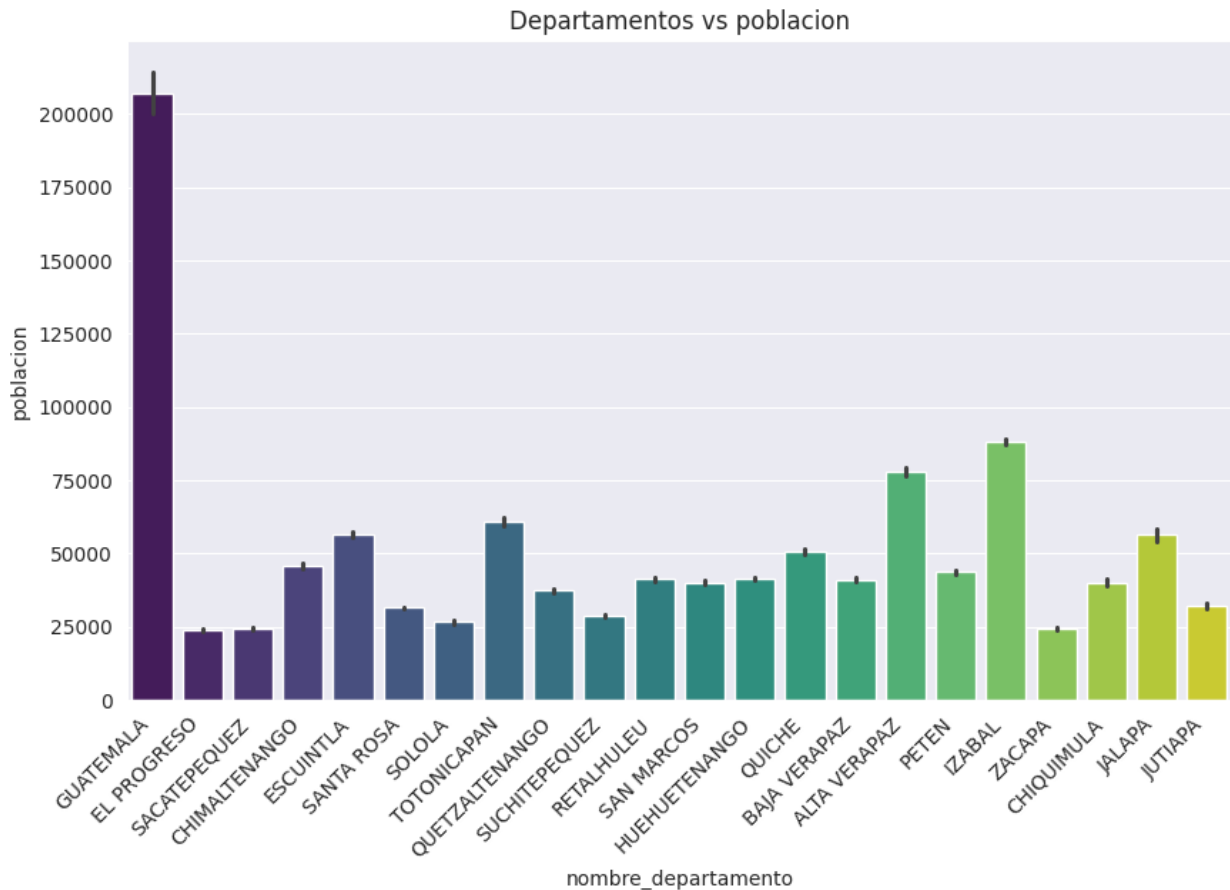
Mapa de calor entre municipios y poblacion

```
plt.figure(figsize=(10, 6))
ax = sns.heatmap(df_muertes_municipio[['poblacion',
'codigo_municipio']].corr(), annot=True)
ax.set_title('Mapa de calor entre municipios y poblacion')
plt.show()
```

Grafica de barras departamentos vs poblacion

```
plt.figure(figsize=(10, 6))
ax = sns.barplot(x='nombre_departamento', y='poblacion',
data=df_muertes_municipio, palette='viridis')
ax.set_xticklabels(ax.get_xticklabels(), rotation=45, ha="right")
ax.set_title('Departamentos vs poblacion')
plt.show()
```



Analisis sobre grafica de barras entre departamentos y poblacion

Se puede observar que el departamento de Guatemala es el que tiene mayor poblacion.

Mapa de calor entre departamentos y poblacion

```
plt.figure(figsize=(10, 6))
ax = sns.heatmap(df_muertes_municipio[['poblacion',
'codigo_departamento']].corr(), annot=True)
ax.set_title('Mapa de calor entre departamentos y poblacion')
plt.show()
```



Conclusiones Generales

1. Tomando en cuenta los resultados obtenidos en los analisis se puede concluir que el departamento de Guatemala es el que tiene mayor cantidad de muertes, esto es porque es el departamento con mas poblacion. Por lo que se debe tomar en cuenta al momento de estar en una sanitaria similar, para poder tomar medidas mas estrictas en este departamento.
2. Se observo que las muertes van aumentando con el tiempo por lo que se debe tomar en cuenta que las medidas sanitarias deben ser mas estrictas al inicio de una pandemia para tratar de mantener las muertes al minimo.
3. Las muertes acumuladas estan restrictamente relacionadas a las muertes nuevas por lo que si se logra controlar las muertes nuevas se lograra controlar el aumento de las muertes acumuladas.