

MUERTES MATERNAS

José Emanuel

Rodríguez Fitta

CARGA DE DATOS

```
muertes_maternas_data = pd.read_csv('/content/Muertes_Maternas.csv', encoding = 'latin')
muertes_maternas_data.head(5)
```

	ANIO_NACIMIENTO	MES_NACIMIENTO	MES_NACIMIENTOD	DIA_NACIMIENTO	EDAD	ESTADO_CONYUGAL	ESTADO_CONYUGALD	ENTIDAD_RESIDENCIA	ENTIDAD_RESIDENCIAD	MUNICIPIO_RESIDENCIA	...	DIA_REG
0	0	0	NO ESPECIFICADO	0	35	4	UNION LIBRE	8	CHIHUAHUA	29	...	
1	0	0	NO ESPECIFICADO	0	20	5	CASADO	8	CHIHUAHUA	47	...	
2	1962	7	JULIO	12	39	5	CASADO	12	GUERRERO	20	...	
3	1974	2	FEBRERO	10	27	5	CASADO	12	GUERRERO	41	...	
4	1980	4	ABRIL	18	20	4	UNION LIBRE	18	NAYARIT	9	...	

5 rows × 58 columns

TRANSFORMACIÓN

Se cambió el valor “MÉXICO” de la columna por el valor “ESTADO DE MÉXICO” de la columna “ENTIDAD_OCURRENCIAD”, para la correcta interpretación en Power BI.

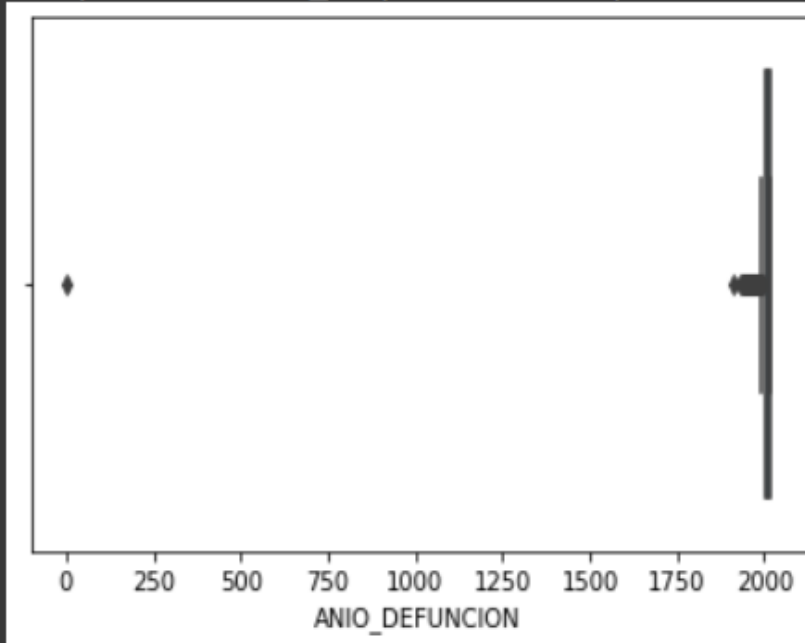
```
# Cambiar Mexico por estado de mexico
muertes_maternas_data['ENTIDAD_OCURRENCIAD'] = muertes_maternas_data['ENTIDAD_OCURRENCIAD'].apply(lambda x: 'ESTADO DE MÉXICO' if x == 'MÉXICO' else x)
```

TRANSFORMACIÓN

Se creo un Data Frame auxiliar con los años de defunción con valor distinto de cero.

```
sns.boxplot(x = 'ANIO_DEFUNCION', data = muertes_maternas_data)
```

```
<matplotlib.axes._subplots.AxesSubplot at 0x7f61c40ed850>
```

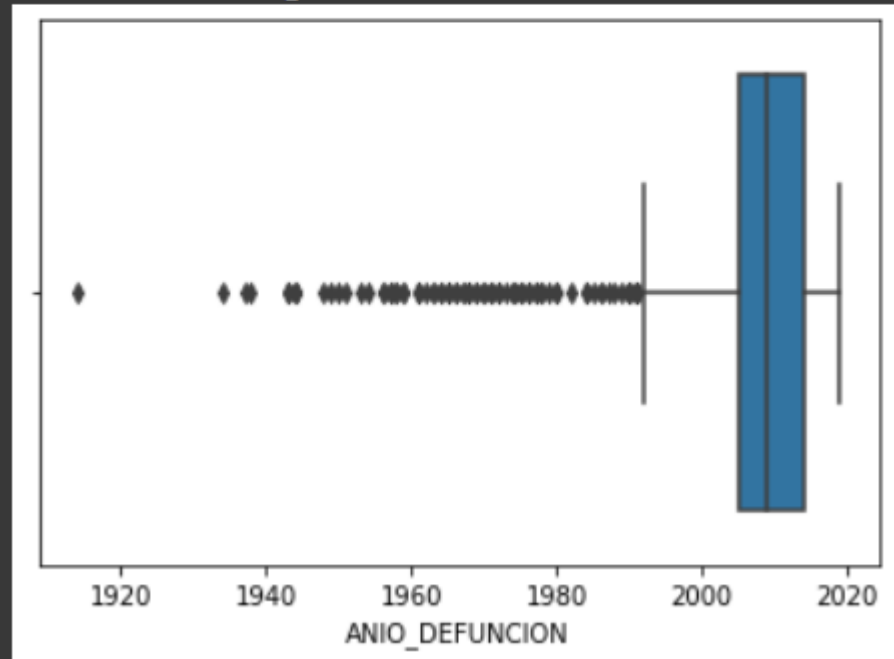


```
muertes_per_anio = muertes_maternas_data[ muertes_maternas_data['ANIO_DEFUNCION'] != 0 ]
```

TRANSFORMACIÓN

```
sns.boxplot(muertes_per_anio['ANIO_DEFUNCION'])  
plt.show()
```

/usr/local/lib/python3.7/dist-packages/seaborn/
FutureWarning



Observamos que la gran mayoría de los registros son a partir del 2002. Por lo que se realizará un análisis para los años superiores a este.

```
muertes_desde_2002 = muertes_per_anio[muertes_per_anio['ANIO_DEFUNCION'] >= 2002]
```

TRANSFORMACIÓN

```
import string
from sklearn.feature_extraction.text import CountVectorizer

# Extracción de las causas de muertes maternas
causa_df = muertes_desde_2002['CAUSA_CIE_4D'].unique()
causa_df = pd.DataFrame(causa_df).rename( columns = {0: 'CAUSA_CIE_4D'})

# Tokenización
punctuation = set(string.punctuation)
def tokenize(sentence):
    tokens = []
    for token in sentence.split():
        new_token = []
        for character in token:
            if character not in punctuation:
                new_token.append(character.lower())
        if new_token:
            tokens.append("".join(new_token))
    return tokens[1:]

# Aplicamos la función tokenize en cada registro de la columna 'CAUSA_CIE_4D'
causa_df['CAUSA_CIE_4D'].apply(tokenize)

demo_vectorizer = CountVectorizer(
    tokenizer = tokenize,
    binary = True
)

real_vectorizer = CountVectorizer(tokenizer = tokenize)
train_X = real_vectorizer.fit_transform(causa_df['CAUSA_CIE_4D'])
```

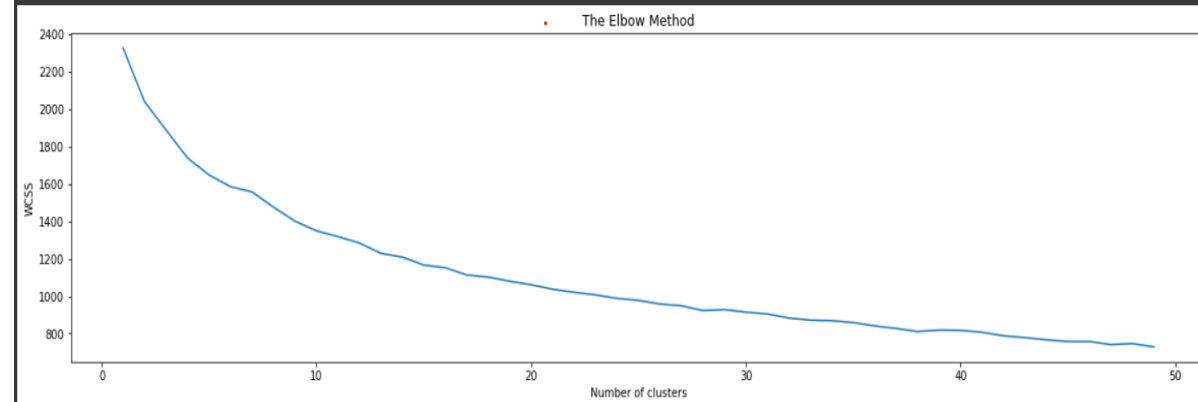
Debido a que se tienen cerca de 200 causas de muerte, se aplicó el algoritmo KMeans para obtener categorías de las causas de muerte que agrupen a las más parecidas. Para esto, lo primero fue convertir a vector cada registro de la columna “CAUSA_CIE_4D”.

TRANSFORMACIÓN

Mediante el atributo `inertia_` podemos observar que el posible número óptimo de clusters es 20, es decir, tomaremos 20 categorías de causas de muertes.

```
from sklearn.cluster import KMeans
wcss = []
for i in range(1, 50):
    kmeans = KMeans(n_clusters = i, init = 'k-means++', random_state = 42)
    kmeans.fit(train_X)
    wcss.append(kmeans.inertia_)
```

```
# Grafica de la suma de las distancias
plt.figure(figsize = (20, 5))
plt.plot(range(1, 50), wcss)
plt.title('The Elbow Method')
plt.xlabel('Number of clusters')
plt.ylabel('WCSS')
plt.show()
```



TRANSFORMACIÓN

Con esto es posible agregar una columna al data frame que clasifique cada registro en una de estas veinte categorías.

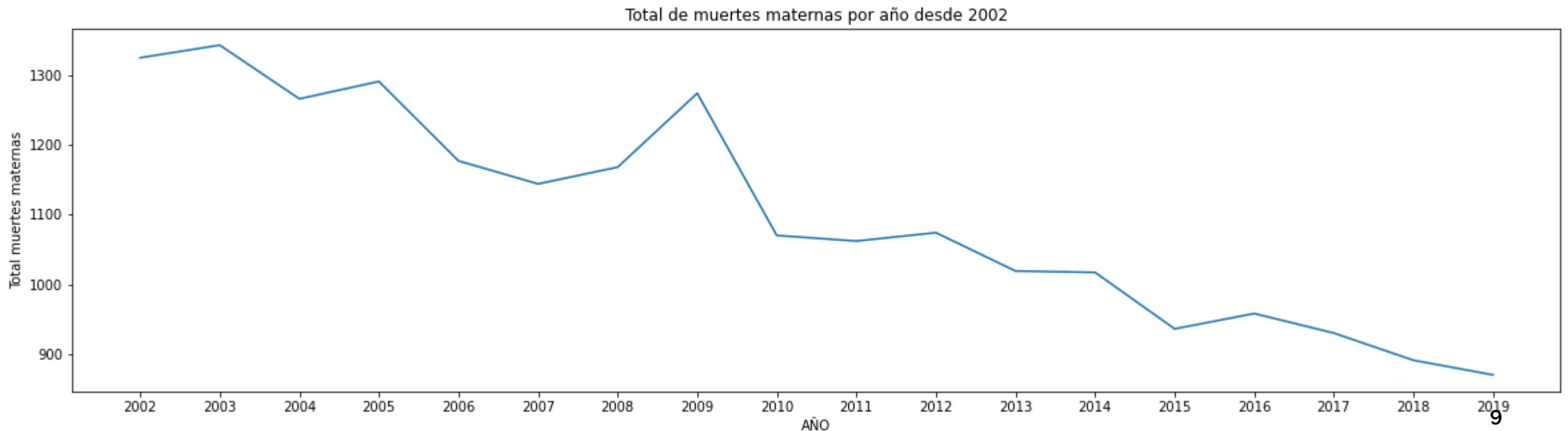
```
muertes_desde_2002 = pd.merge(muertes_desde_2002, causa_df, on = 'CAUSA_CIE_4D')
```

```
muertes_desde_2002.head(5)
```

ICACION	MES_CERTIFICACION	MES_CERTIFICACIOND	DIA_CERTIFICACION	ANIO_BASE_DATOS	RAZON_MORTALIDAD_MATERNA	RAZON_MORTALIDAD_MATERNAD	EDAD_QUINQUENAL	EDAD_QUINQUENALD	clasificacion
0	0	NO ESPECIFICADO	0	2002	1	MUERTES MATERNAS PARA LA RAZÓN DE MORTALIDAD M...	4	25 a 29 años	0
0	0	NO ESPECIFICADO	0	2002	1	MUERTES MATERNAS PARA LA RAZÓN DE MORTALIDAD M...	3	20 a 24 años	0
0	0	NO ESPECIFICADO	0	2002	1	MUERTES MATERNAS PARA LA RAZÓN DE MORTALIDAD M...	3	20 a 24 años	0
0	0	NO ESPECIFICADO	0	2002	1	MUERTES MATERNAS PARA LA RAZÓN DE MORTALIDAD M...	3	20 a 24 años	0
0	0	NO ESPECIFICADO	0	2002	1	MUERTES MATERNAS PARA LA RAZÓN DE MORTALIDAD M...	4	25 a 29 años	0

VISUALIZACIONES

Se observa una importante disminución en el total de muertes maternas por año del 2002 al 2019

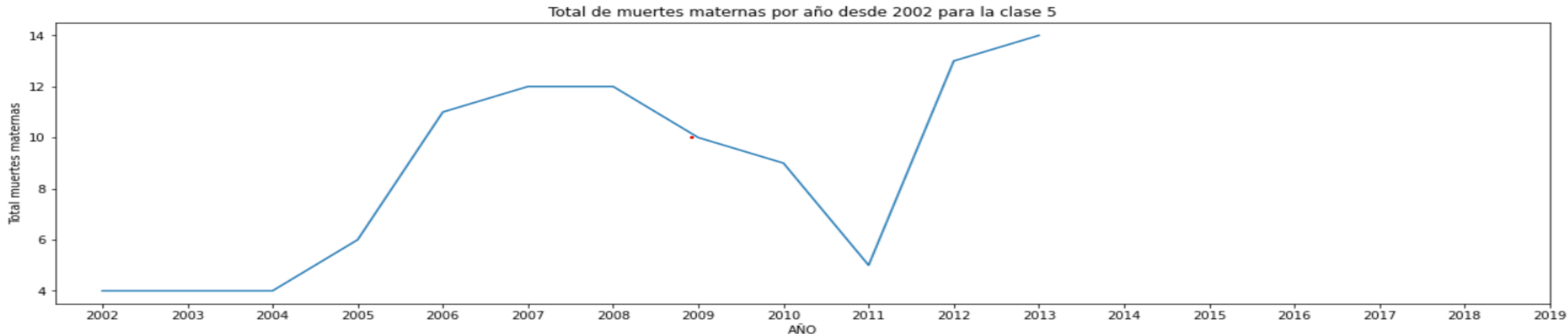


VISUALIZACIONES

La clase 5 corresponde a las muertes relacionadas con el VIH. A pesar de existir una tendencia creciente en el total de muertes por año, vemos que en el 2013, se dejaron de tener registros de muertes dentro de esta categoría.

```
causa_df[causa_df['clasificacion'] == 5]['CAUSA_CIE_4D'].to_list()
```

```
['B238 ENFERMEDAD POR VIH, RESULTANTE EN OTRAS AFECCIONES ESPECIFICADAS',  
'B206 ENFERMEDAD POR VIH, RESULTANTE EN NEUMONÍA POR PNEUMOCYSTIS CARINII',  
'B208 ENFERMEDAD POR VIH, RESULTANTE EN OTRAS ENFERMEDADES INFECCIOSAS O PARASITARIAS',  
'B207 ENFERMEDAD POR VIH, RESULTANTE EN INFECCIONES MÚLTIPLES',  
'B227 ENFERMEDAD POR VIH, RESULTANTE EN ENFERMEDADES MÚLTIPLES CLASIFICADAS EN OTRA PARTE',  
'B201 ENFERMEDAD POR VIH, RESULTANTE EN OTRAS INFECCIONES BACTERIANAS',  
'B200 ENFERMEDAD POR VIH, RESULTANTE EN INFECCIÓN POR MICOBACTERIAS',  
'B220 ENFERMEDAD POR VIH, RESULTANTE EN ENCEFALOPATÍA',  
'B209 ENFERMEDAD POR VIH, RESULTANTE EN ENFERMEDAD INFECCIOSA O PARASITARIA NO ESPECIFICADA',  
'B232 ENFERMEDAD POR VIH, RESULTANTE EN ANORMALIDADES INMUNOLÓGICAS Y HEMATOLÓGICAS, NO CLASIFICADAS EN OTRA PARTE',  
'B222 ENFERMEDAD POR VIH, RESULTANTE EN SÍNDROME CAQUÉTICO',  
'B205 ENFERMEDAD POR VIH, RESULTANTE EN OTRAS MICOSIS',  
'B203 ENFERMEDAD POR VIH, RESULTANTE EN OTRAS INFECCIONES VIRALES',  
'B204 ENFERMEDAD POR VIH, RESULTANTE EN CANDIDIASIS',  
'B218 ENFERMEDAD POR VIH, RESULTANTE EN OTROS TUMORES MALIGNOS']
```

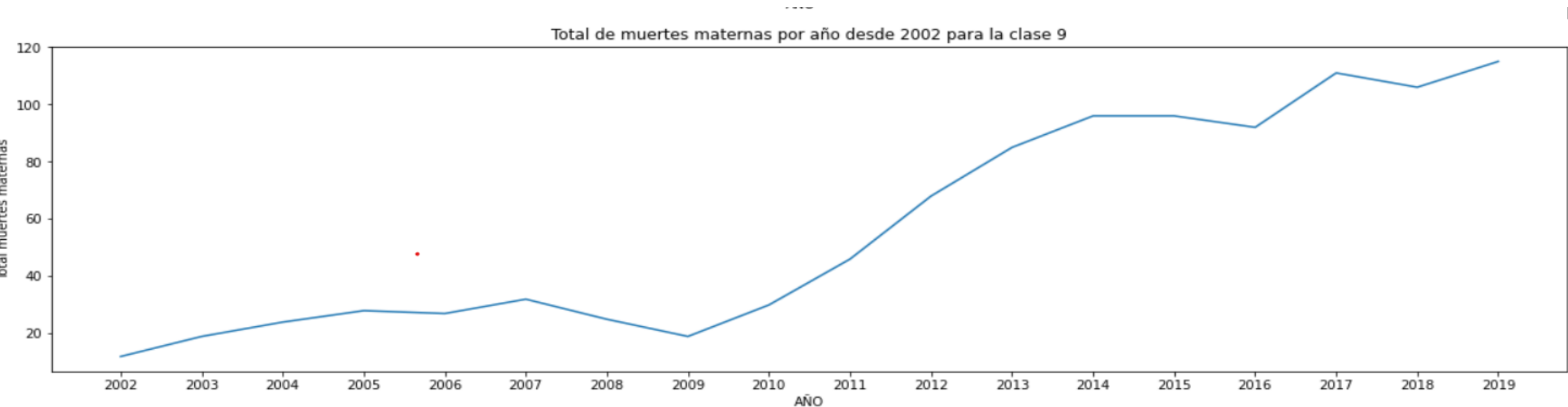


VISUALIZACIONES

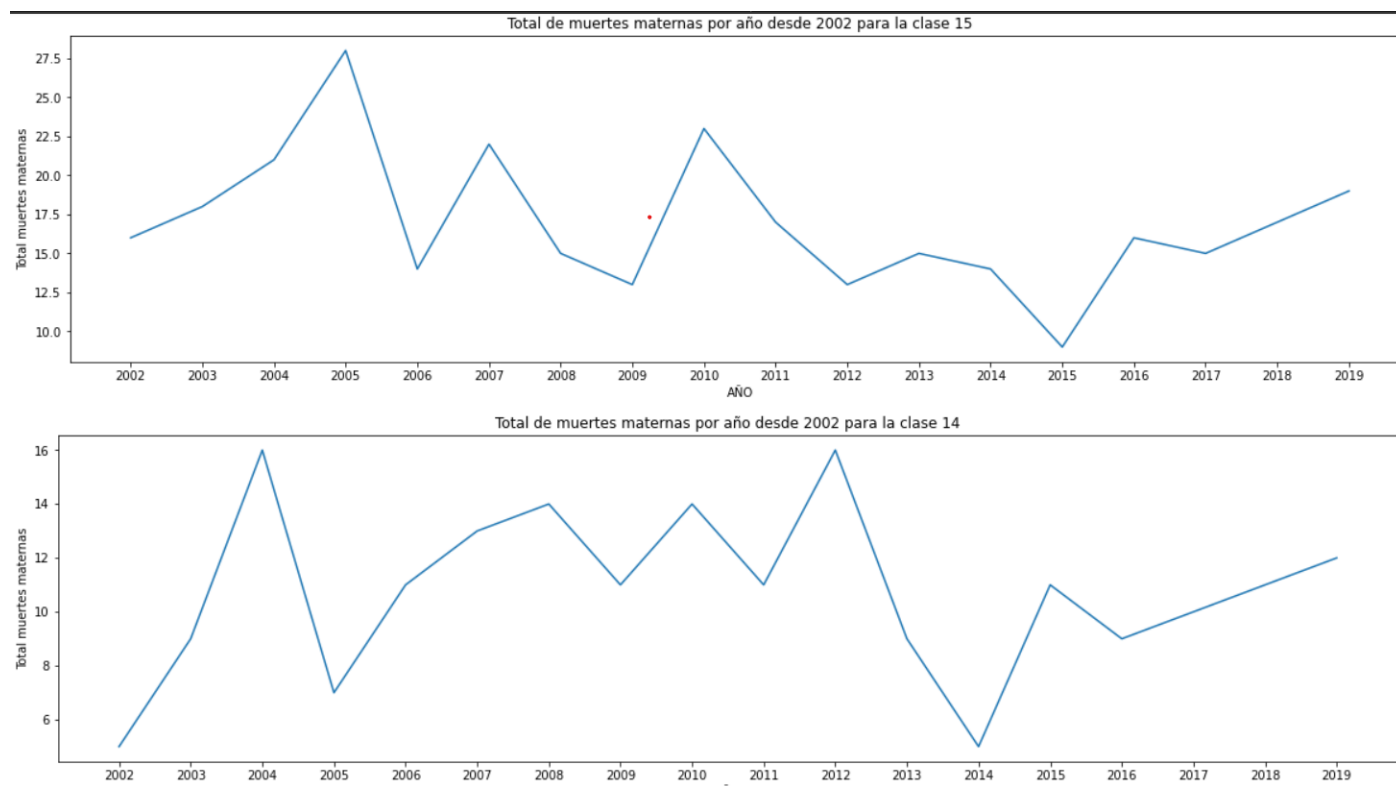
Podemos observar que existe una clara y preocupante tendencia creciente para el total de muertes por año de la clase 9, que corresponde a muertes maternas ocurridas 42 días después del parto y antes del año.

```
causa_df[causa_df['clasificacion'] == 9]['CAUSA_CIE_4D'].to_list()
```

```
['096X MUERTE MATERNA DEBIDA A CUALQUIER CAUSA OBSTÉTRICA QUE OCURRE DESPUÉS DE 42 DÍAS PERO ANTES DE UN AÑO DEL PARTO',  
'0961 MUERTE POR CAUSA OBSTÉTRICA INDIRECTA QUE OCURRE DESPUÉS DE 42 DÍAS PERO ANTES DE UN AÑO DEL PARTO',  
'0960 MUERTE POR CAUSA OBSTÉTRICA DIRECTA QUE OCURRE DESPUÉS DE 42 DÍAS PERO ANTES DE UN AÑO DEL PARTO',  
'0969 MUERTE POR CAUSA OBSTÉTRICA NO ESPECIFICADA, QUE OCURRE DESPUÉS DE 42 DÍAS PERO ANTES DE UN AÑO DEL PARTO']
```



VISUALIZACIONES



Para las clases 14 y 15 podemos observar que existe una tendencia alcista a partir del año 2015. Ambas clases tienen que ver con muertes relacionadas con el aborto. En un caso debido a complicaciones por infección y en el otro debido a complicaciones por hemorragias. Deberían tomarse medidas para que esta tendencia se revierta.

```
causa_df[causa_df['clasificacion'] == 15]['CAUSA_CIE_4D'].to_list()

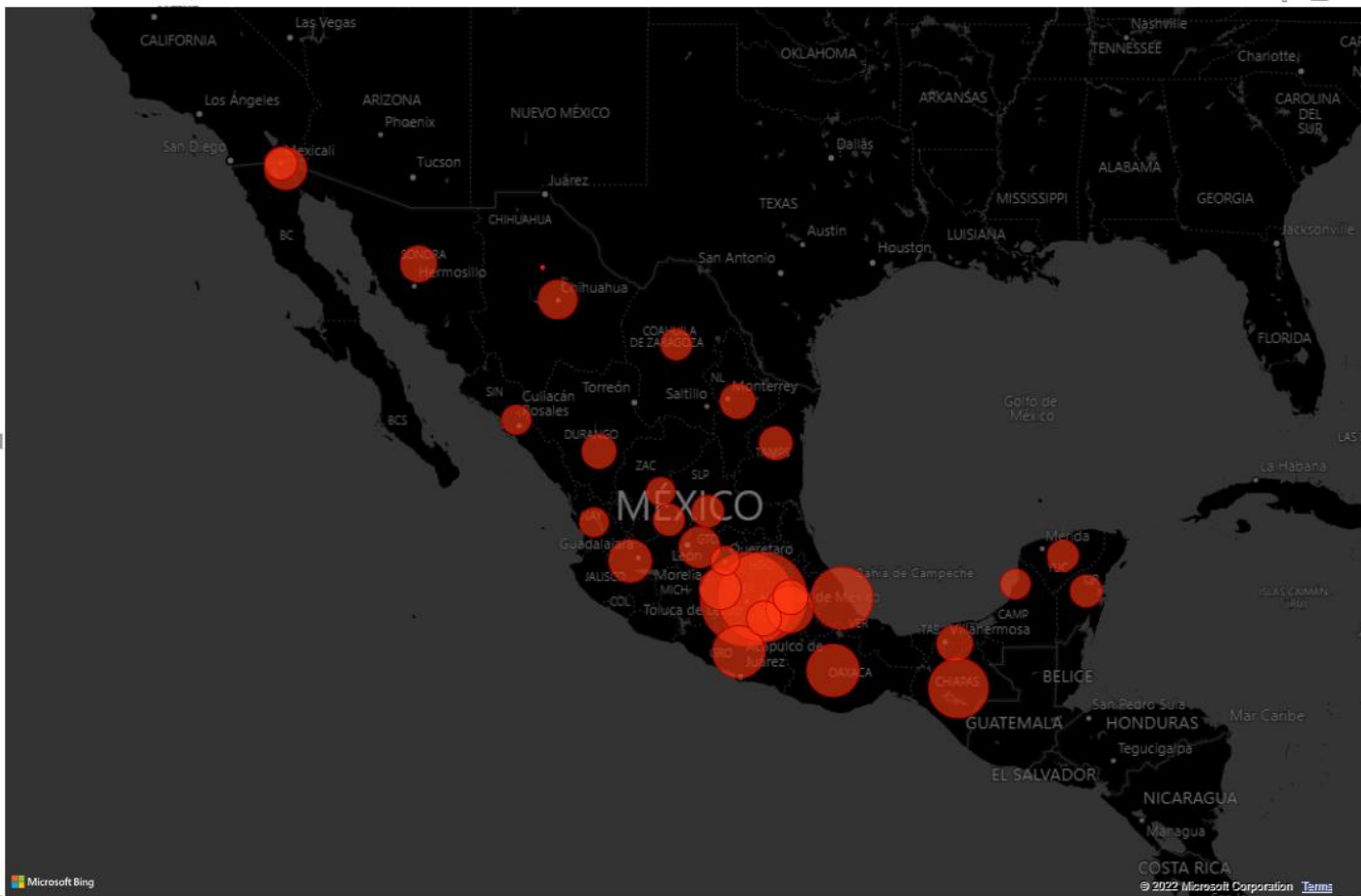
['0065 ABORTO NO ESPECIFICADO COMPLETO O NO ESPECIFICADO, COMPLICADO CON INFECCIÓN GENITAL Y PELVIANA',
'0060 ABORTO NO ESPECIFICADO INCOMPLETO, COMPLICADO CON INFECCIÓN GENITAL Y PELVIANA',
'0035 ABORTO ESPONTÁNEO COMPLETO O NO ESPECIFICADO, COMPLICADO CON INFECCIÓN GENITAL Y PELVIANA',
'0050 OTRO ABORTO INCOMPLETO, COMPLICADO CON INFECCIÓN GENITAL Y PELVIANA',
'0030 ABORTO ESPONTÁNEO INCOMPLETO, COMPLICADO CON INFECCIÓN GENITAL Y PELVIANA',
'0055 OTRO ABORTO COMPLETO O NO ESPECIFICADO, COMPLICADO CON INFECCIÓN GENITAL Y PELVIANA',
'0040 ABORTO MÉDICO INCOMPLETO, COMPLICADO CON INFECCIÓN GENITAL Y PELVIANA',
'0045 ABORTO MÉDICO COMPLETO O NO ESPECIFICADO, COMPLICADO CON INFECCIÓN GENITAL Y PELVIANA']

causa_df[causa_df['clasificacion'] == 14]['CAUSA_CIE_4D'].to_list()

['0066 ABORTO NO ESPECIFICADO COMPLETO O NO ESPECIFICADO, COMPLICADO POR HEMORRAGIA EXCESIVA O TARDÍA',
'0036 ABORTO ESPONTÁNEO COMPLETO O NO ESPECIFICADO, COMPLICADO POR HEMORRAGIA EXCESIVA O TARDÍA',
'0061 ABORTO NO ESPECIFICADO INCOMPLETO, COMPLICADO POR HEMORRAGIA EXCESIVA O TARDÍA',
'0056 OTRO ABORTO COMPLETO O NO ESPECIFICADO, COMPLICADO POR HEMORRAGIA EXCESIVA O TARDÍA',
'0051 OTRO ABORTO INCOMPLETO, COMPLICADO POR HEMORRAGIA EXCESIVA O TARDÍA',
'0031 ABORTO ESPONTÁNEO INCOMPLETO, COMPLICADO POR HEMORRAGIA EXCESIVA O TARDÍA',
'0041 ABORTO MÉDICO INCOMPLETO, COMPLICADO POR HEMORRAGIA EXCESIVA O TARDÍA']
```

VISUALIZACIONES

Muertes maternas por Aborto



Se puede observar que la mayor cantidad de muertes debidas al aborto (infección o hemorragia) se localizan en el centro del país. Por lo cual son estas entidades, quienes deberán tomar medidas para revertir la preocupante tendencia.

VISUALIZACIONES

La mayor cantidad de muertes maternas son de personas que no tienen ningún seguro médico y de poblaciones realmente pequeñas.

Deben reformarse las políticas publicas para poder llevar servicios médicos a dichas comunidades y atender a las personas sin este servicio.

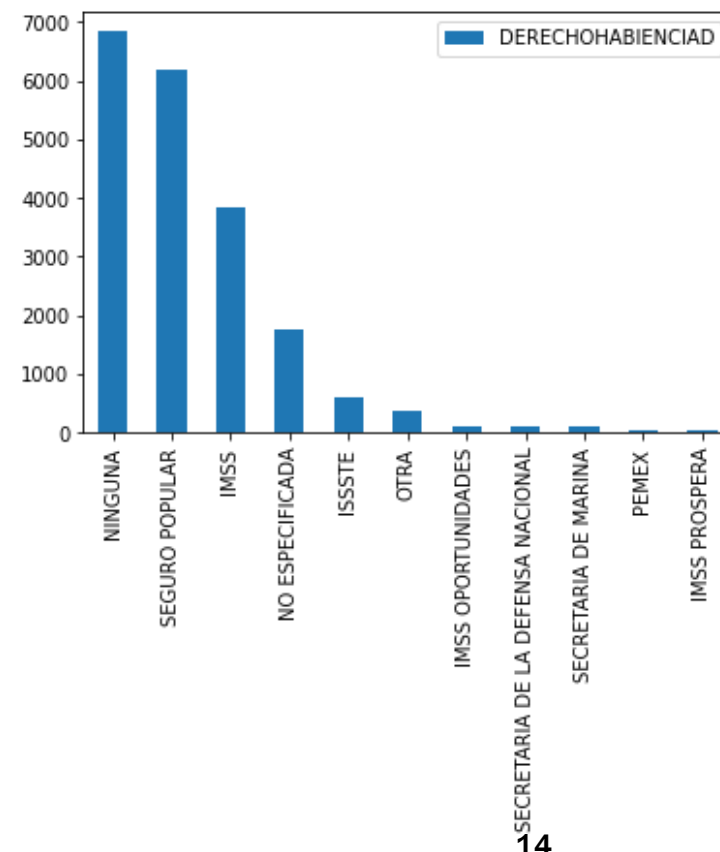
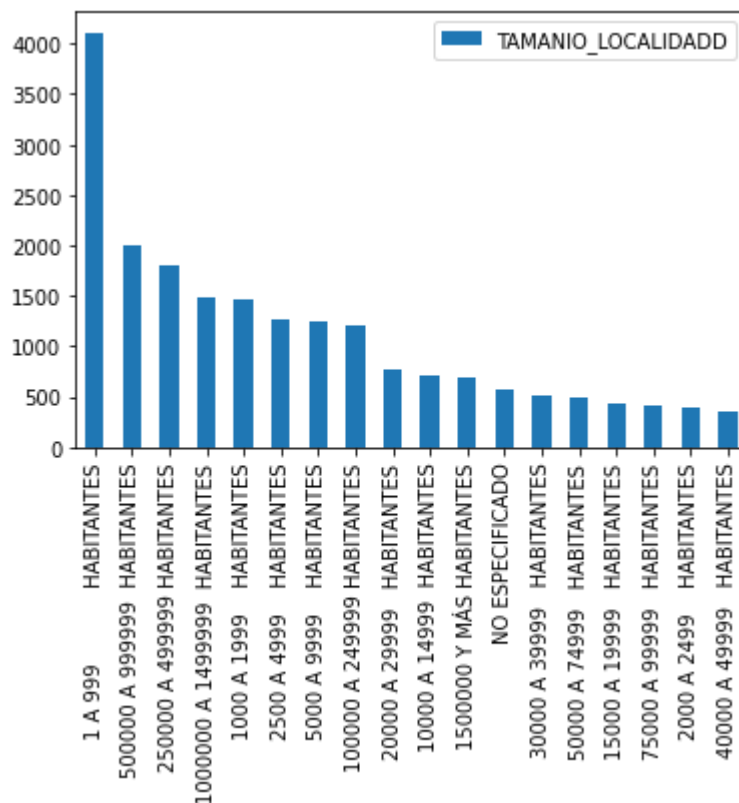
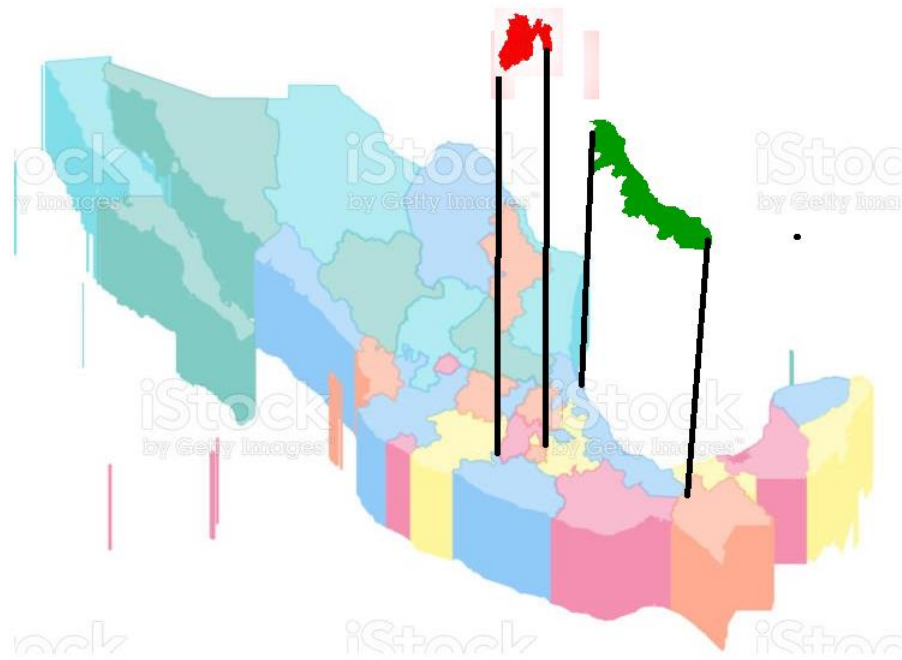


GRÁFICO INTERESANTE



Me es realmente difícil poder dibujar la visualización que me gustaría hacer, pero la pienso más o menos como un gráfico de barras interactivo donde al seleccionar uno o varios estados de la republica puedan resaltar de manera proporcional a la cantidad de muertes maternas que existen y de acuerdo a los filtros proporcionados, año, causa, etc.