



Evolución causas de Muerte en Argentina

Ing. Facundo Laureano
Martinez

Motivación

- Factores incidentes
- Esperanza de vida poblacional
- Avances científicos

Audiencia

- Población
- Profesionales de salud

Preguntas - Hipótesis

- ¿Hay alguna relación entre las muertes causadas en la población Argentina por deficiencia nutricional, VIH, hepatitis aguda y causas hepáticas?
- ¿Cómo podría ser la evolución de los casos de estas enfermedades en Argentina a lo largo de los años respecto a Latinoamérica?
- ¿Puede haber afectado los avances tecnológicos durante a lo largo del tiempo a estas enfermedades?

EDA

El dataset esta compuesto por:

- 7233 filas
- Las columnas mas importantes son: Año, País y el tipo de enfermedad

EDA

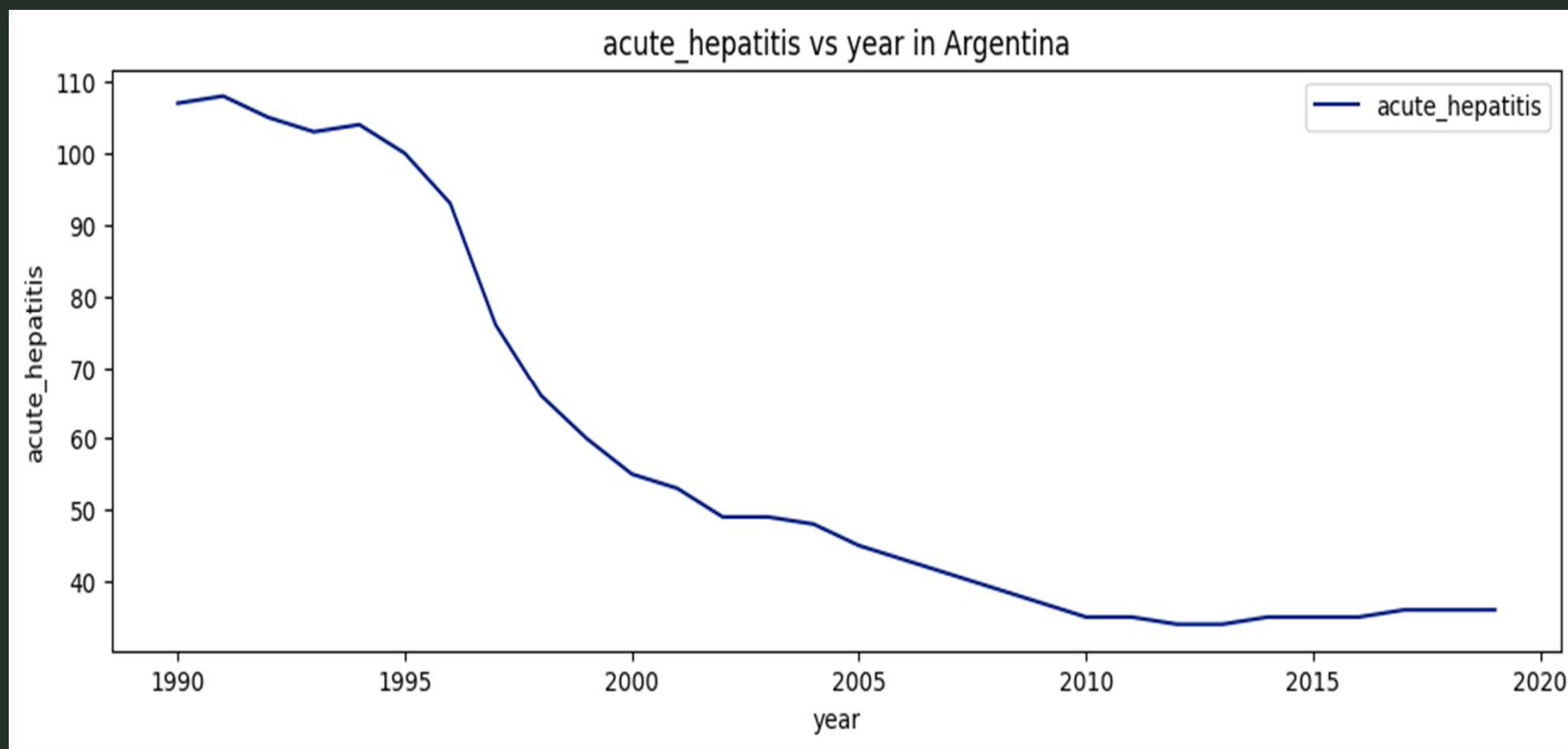
Como se decide trabajar sobre el país Argentina, se filtran los datos para este país y se dropean las columnas que contienen las enfermedades que no se desean evaluar en esta oportunidad.

```
[ ] df.info()
```

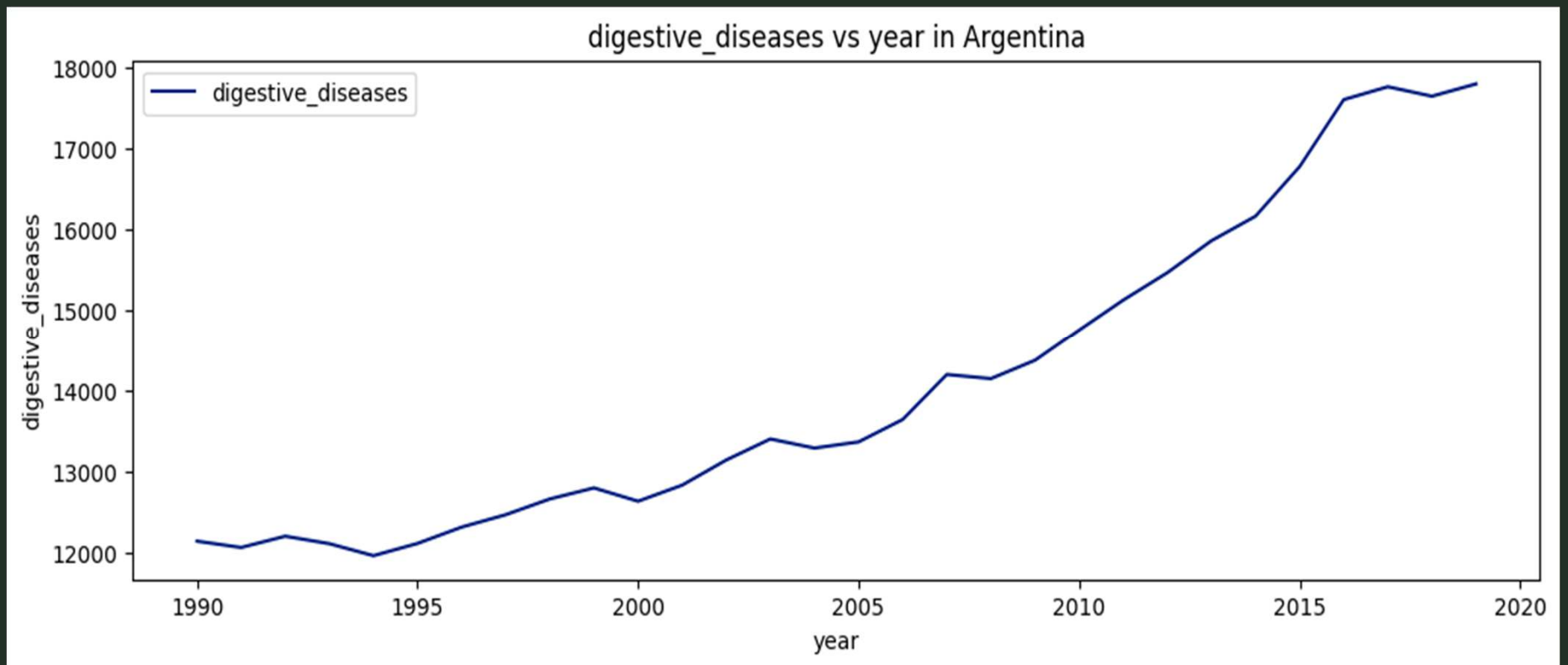
```
<class 'pandas.core.frame.DataFrame'>  
RangeIndex: 7273 entries, 0 to 7272  
Data columns (total 8 columns):  
#   Column                                Non-Null Count  Dtype    
---  ---                                -  
0   country                               7273 non-null   object   
1   year                                 7273 non-null   int64    
2   nutritional_deficiency               6840 non-null   float64  
3   hiv/aids                             6840 non-null   float64  
4   chronic_respiratory_diseases         6840 non-null   float64  
5   chronic_liver_diseases               6840 non-null   float64  
6   digestive_diseases                   6840 non-null   float64  
7   acute_hepatitis                      6840 non-null   float64  
dtypes: float64(6), int64(1), object(1)  
memory usage: 454.7+ KB
```

```
[ ] df_filtered=df[(df['country']=='Argentina')]
```

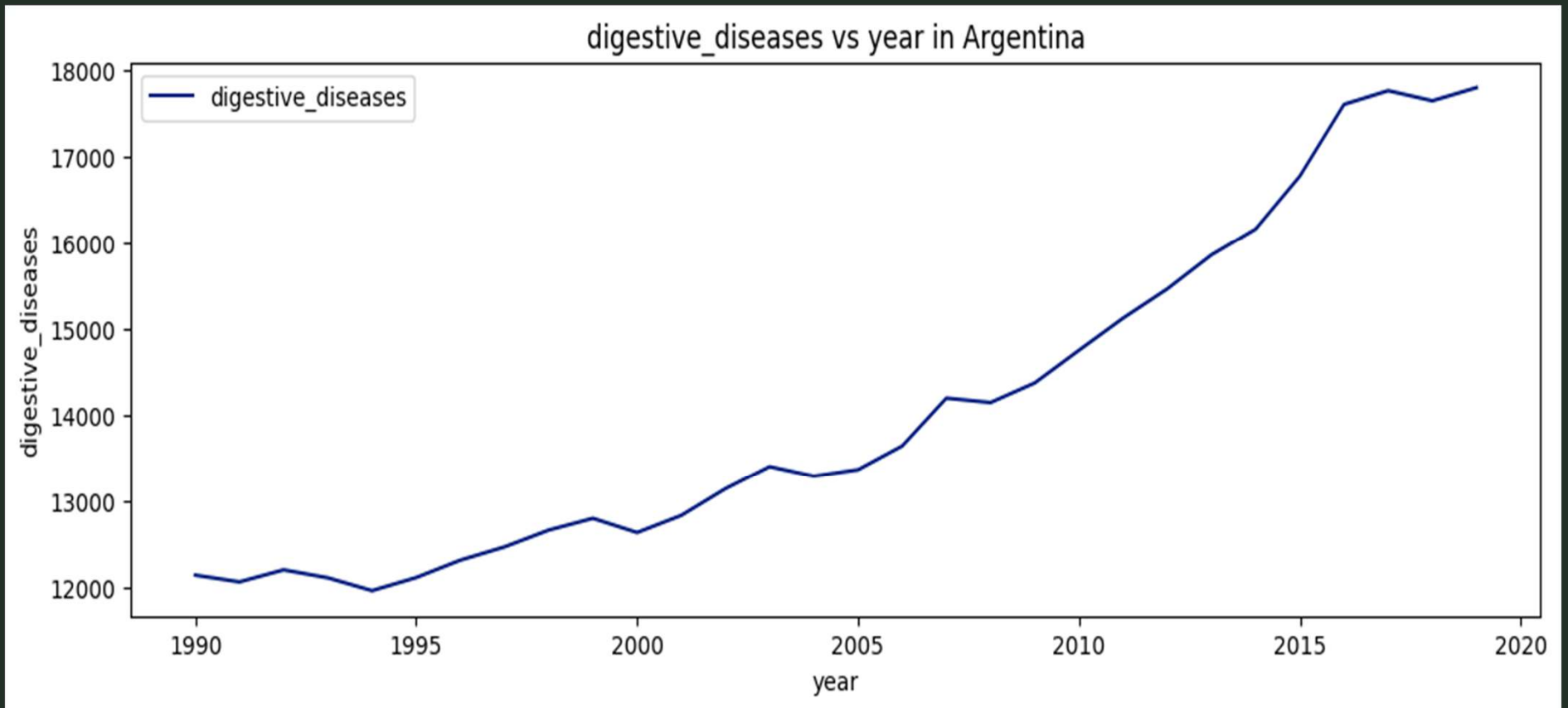
Grafico los casos de las diferentes enfermedades evaluadas en Argentina



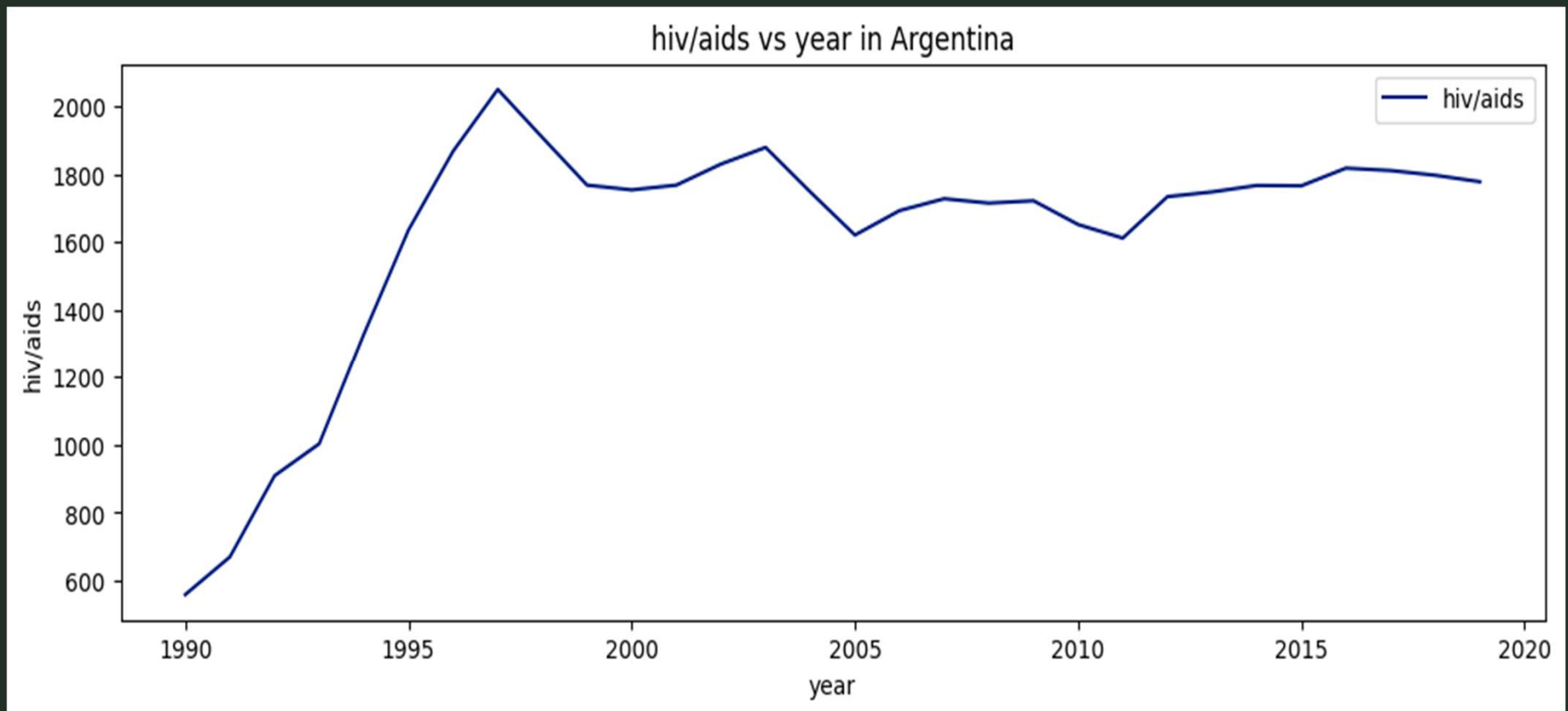
Los casos de hepatitis en la Argentina a lo largo de los años han bajado, lo cual se debe a los avances tecnológicos y de medicina en el tiempo.



Lo mismo sucede con las muertes debido a problemas digestivos, ha habido un aumento significativo

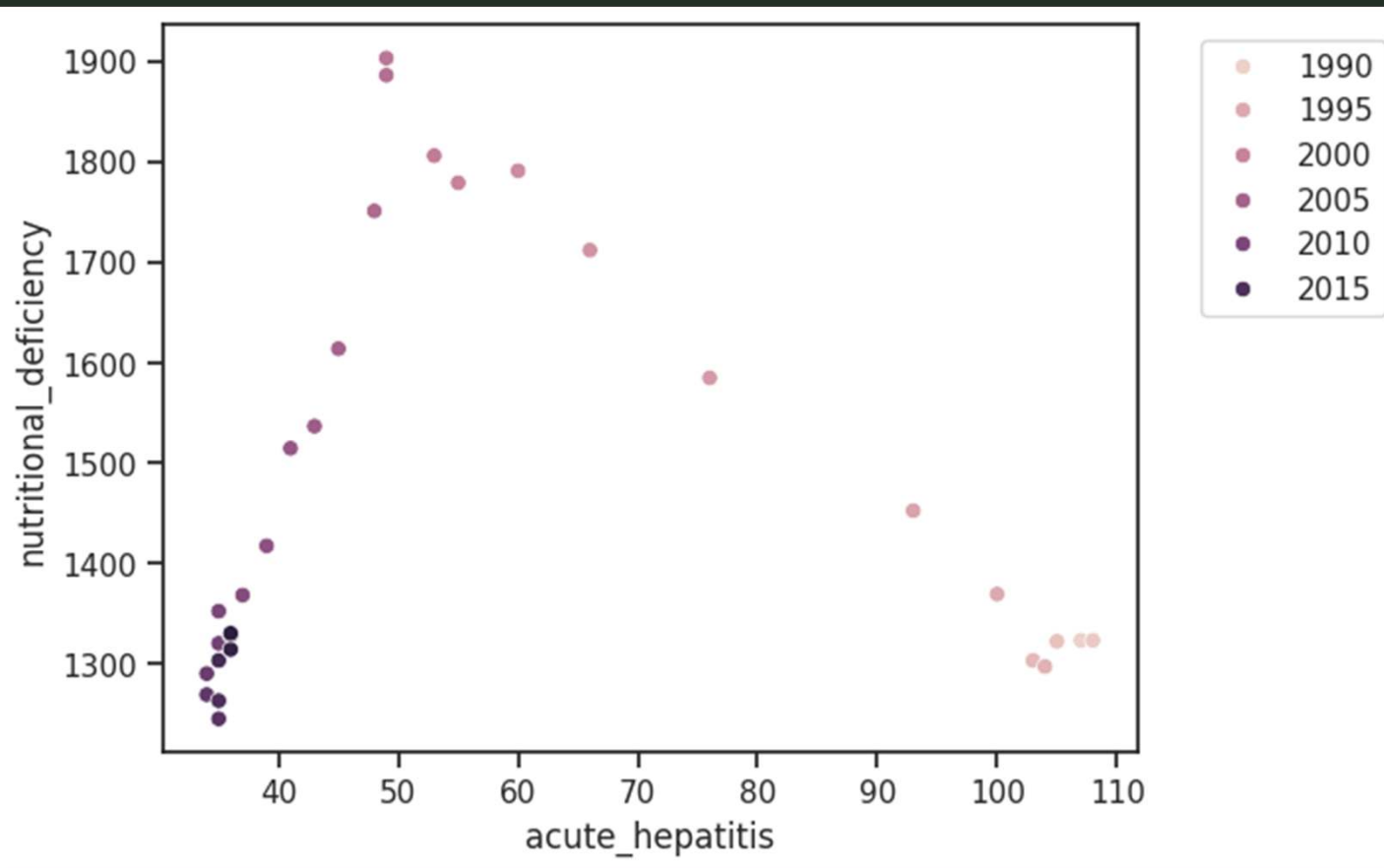


Lo mismo sucede con las muertes debido a problemas digestivos, ha habido un aumento significativo

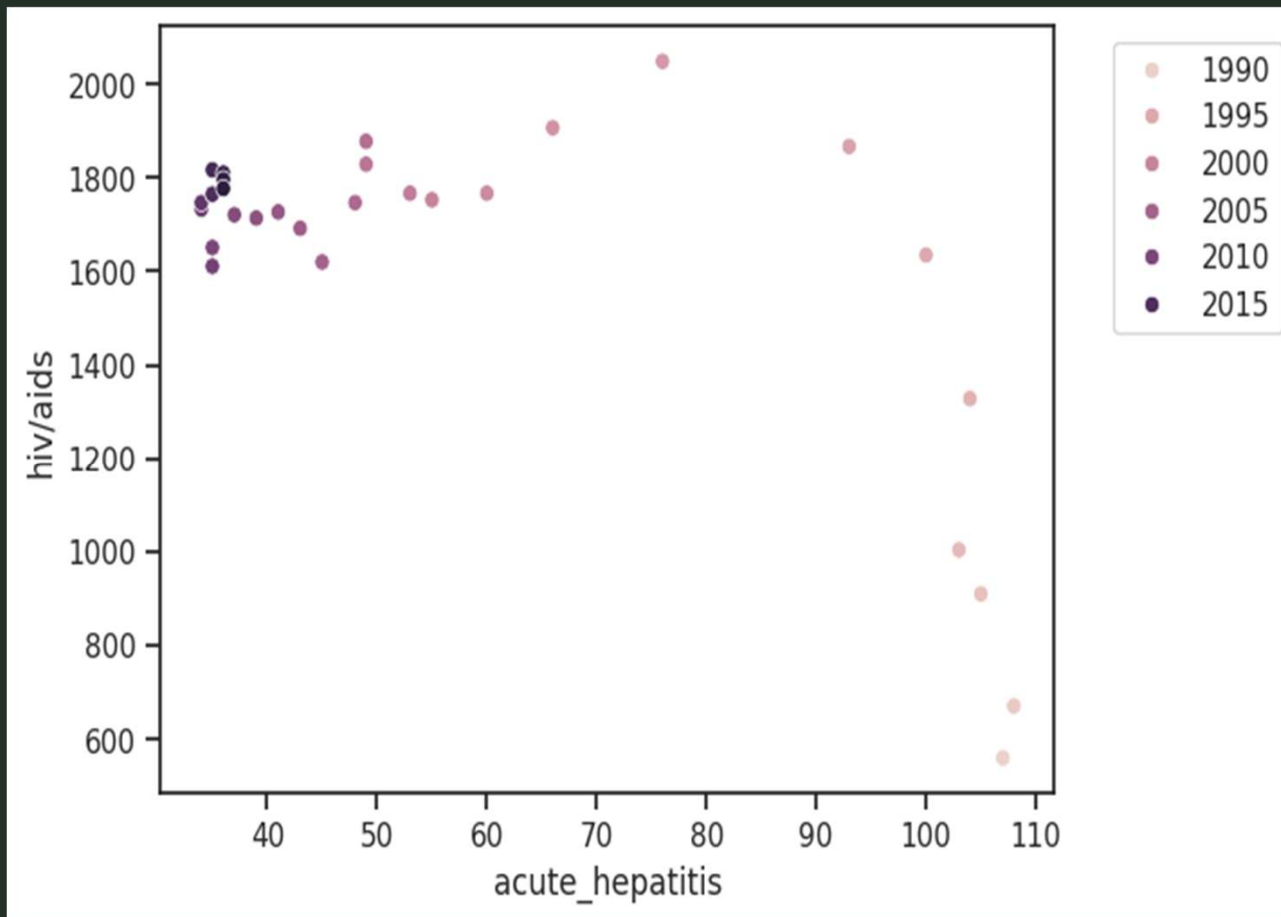


En el caso de las muertes causadas debido vih/sida se observa una pendiente en crecimiento durante la época de los 90s, llegando a un pico máximo alrededor de 1998, luego la tendencia es a estabilizarse. Respetto a esto podríamos ver también cuales fueron los índices de contagios, ya que si bien tiene mucha influencia los nuevos medicamento y la evolución en tecnologías medicas, también seria importante determinar cuantas personas se contagian por año

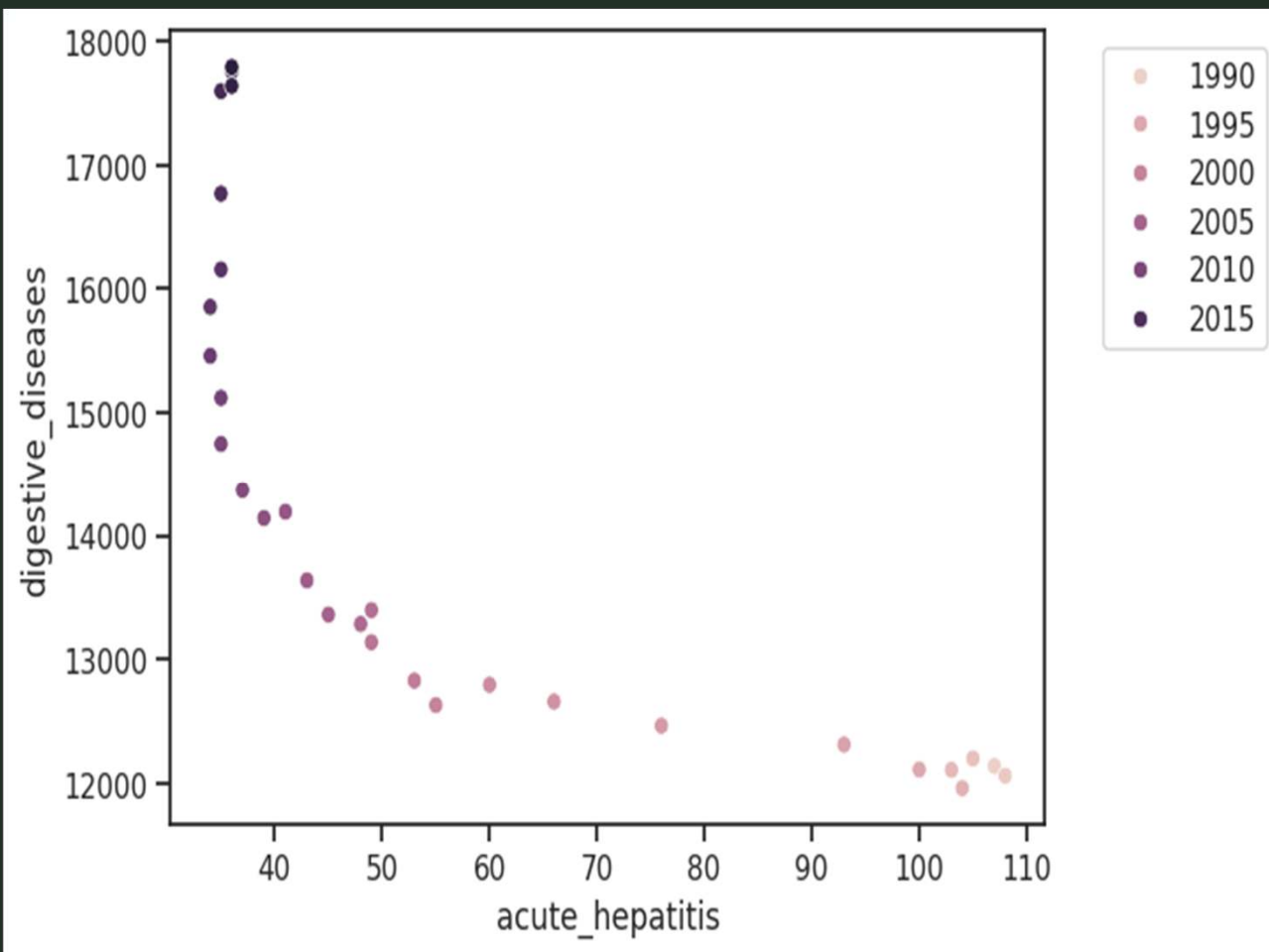
Grafico para comparar dos causas de enfermedades en Argentina



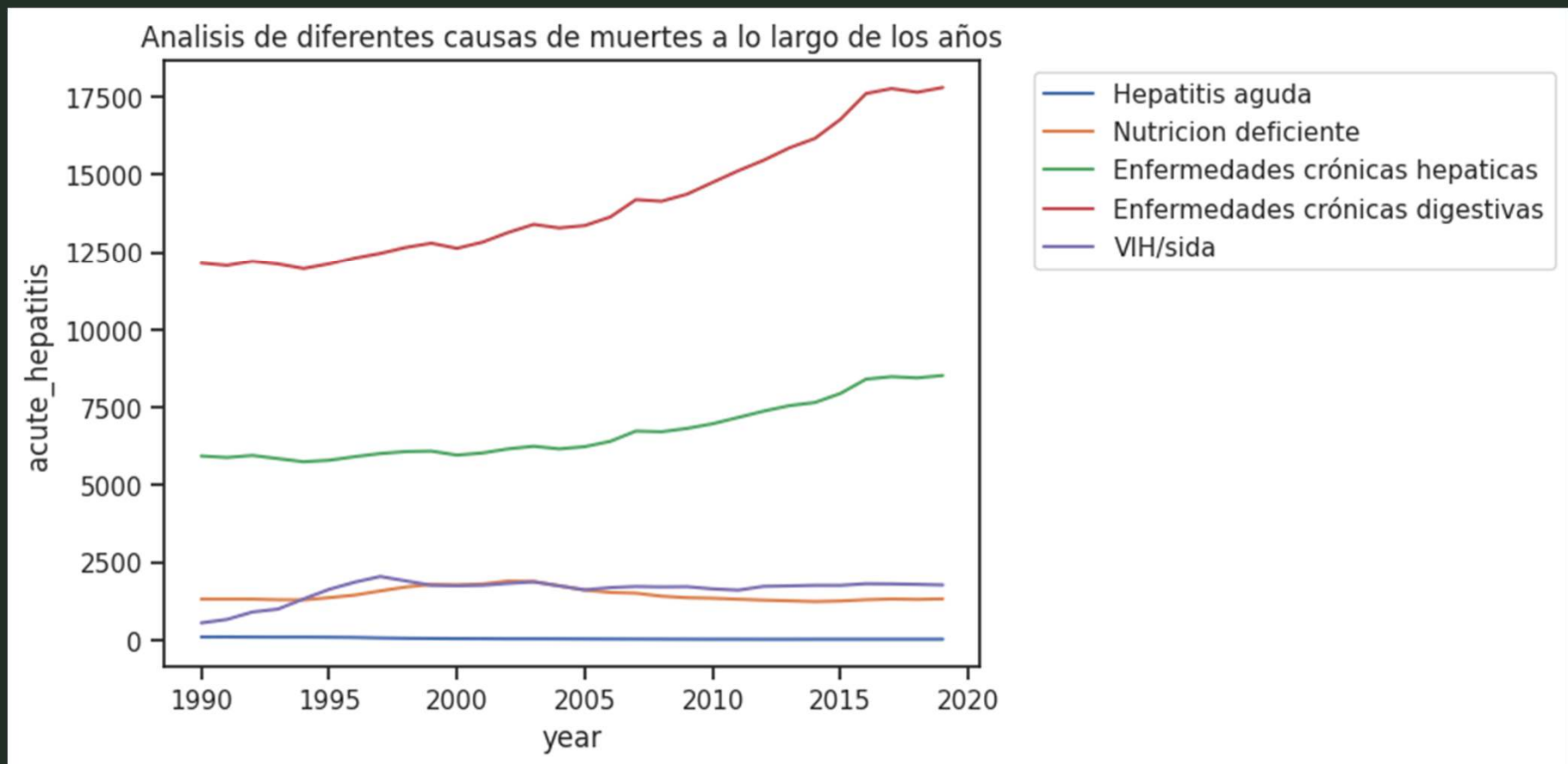
Hay relación entre
Hepatitis aguda y
deficiencia nutricional



No hay relación entre
Hepatitis aguda y VIH



No hay relación entre
Hepatitis aguda y
enfermedades
digestivas



Se grafican todas las enfermedades evaluadas para determinar cual es la que mas aumento a lo largo del tiempo, siendo esta las enfermedades crónicas digestivas, seguidas por enfermedades crónicas hepáticas.

Data Wranling

En este caso se evaluaron los datos nulos y se los reemplazo por 0

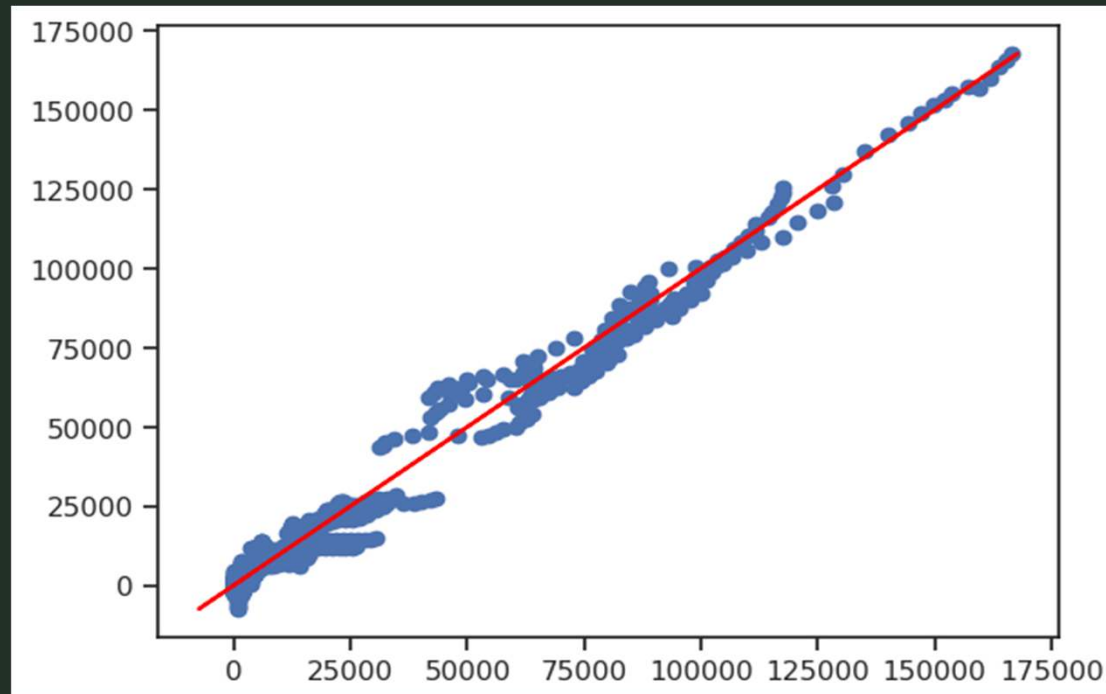
```
[ ] df_no_nulls = df.replace (np.NaN, value=0)  
    df_no_nulls.isnull().sum()
```

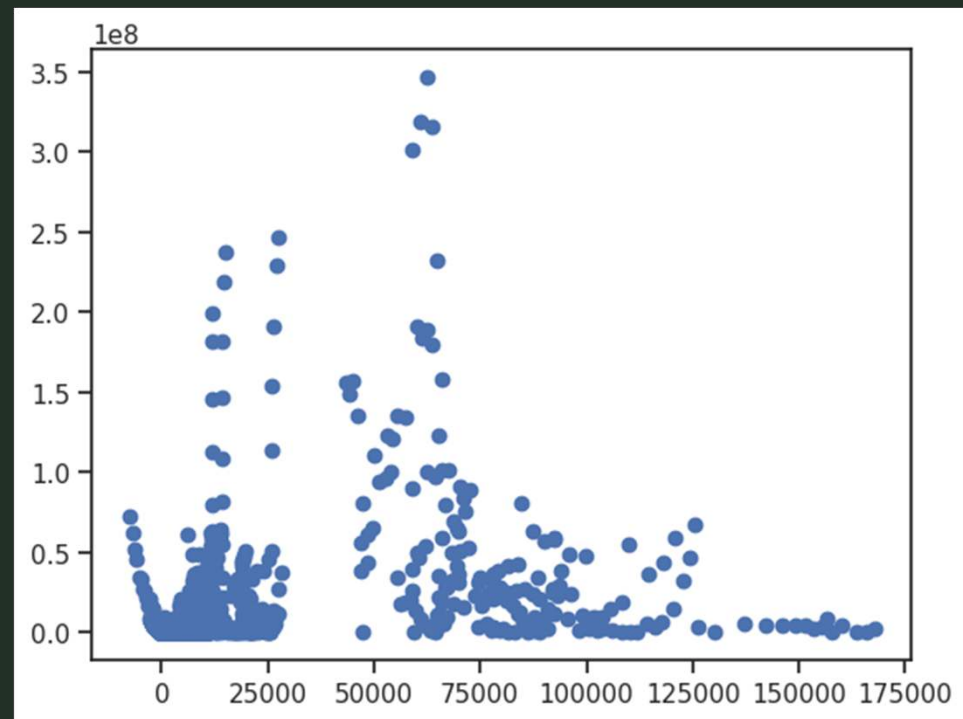
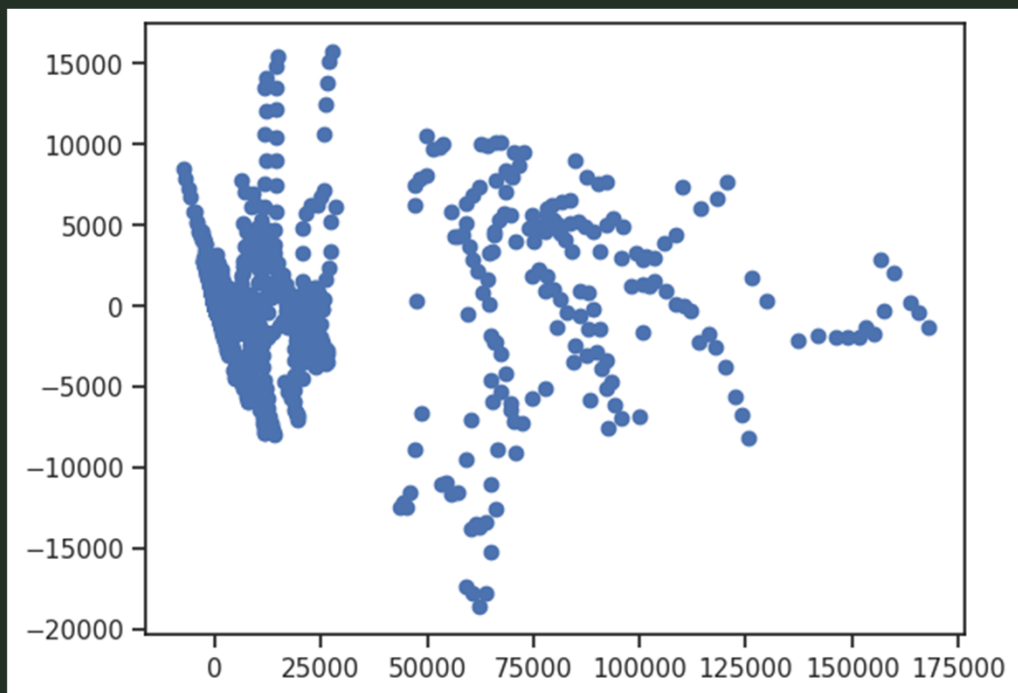
```
country      0  
year         0  
nutritional_deficiency  0  
hiv/aids     0  
chronic_respiratory_diseases  0  
chronic_liver_diseases  0  
digestive_diseases  0  
acute_hepatitis  0  
dtype: int64
```

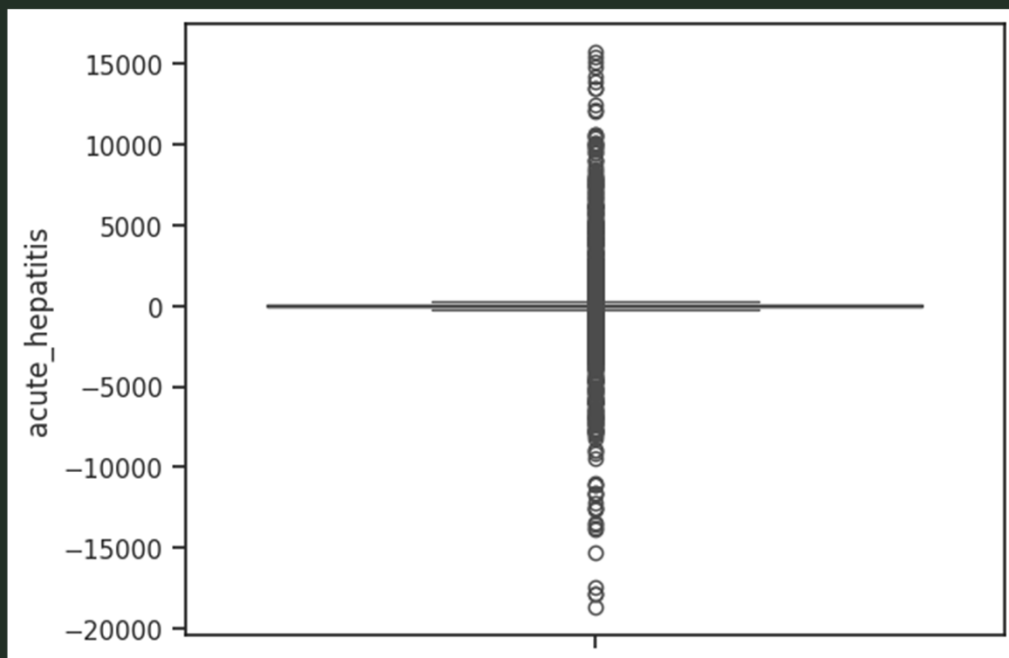
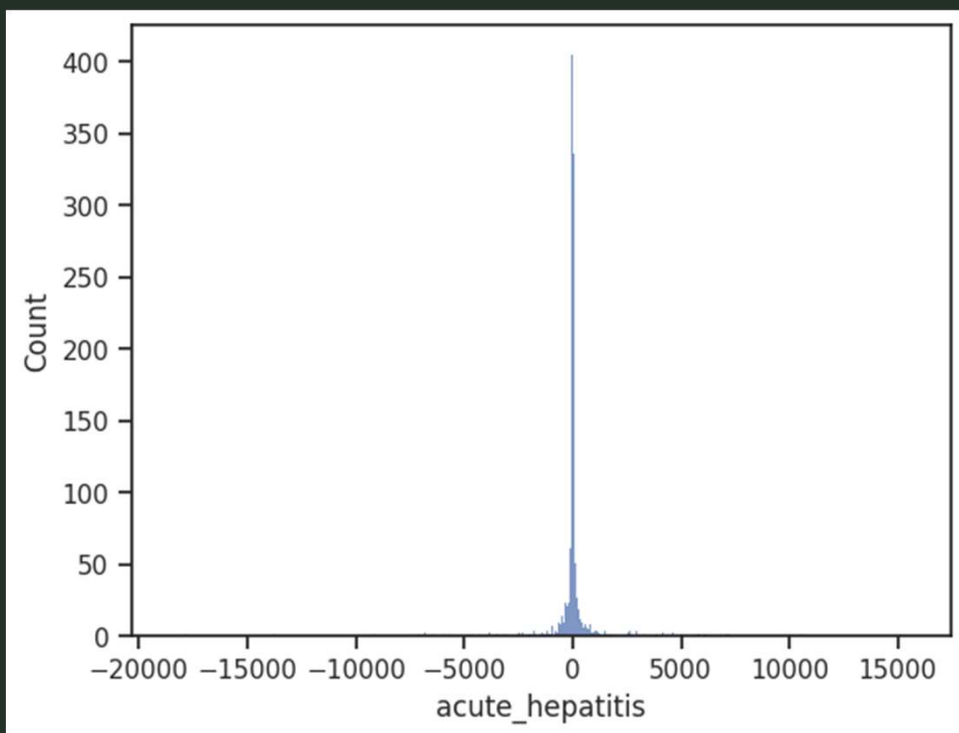
Selección de modelos

Para predecir y graficar tendencias se utilizaron modelos de regresión lineal debido a que las variables son dependientes del tiempo.

Hepatitis aguda







Se entreno y testeo el modelo obteniendo las siguientes métricas:

XGBRegressor

```
MEAN: 145.1706254391184  
MSE: 971969.1662563767  
RMSE: 985.8849660362899  
MAPE: 34342515194569.742
```

LGBMRegressor

```
MEAN: 175.73234126904381  
MSE: 1063355.163918605  
RMSE: 1031.1911384018993  
MAPE: 478096543971845.0
```

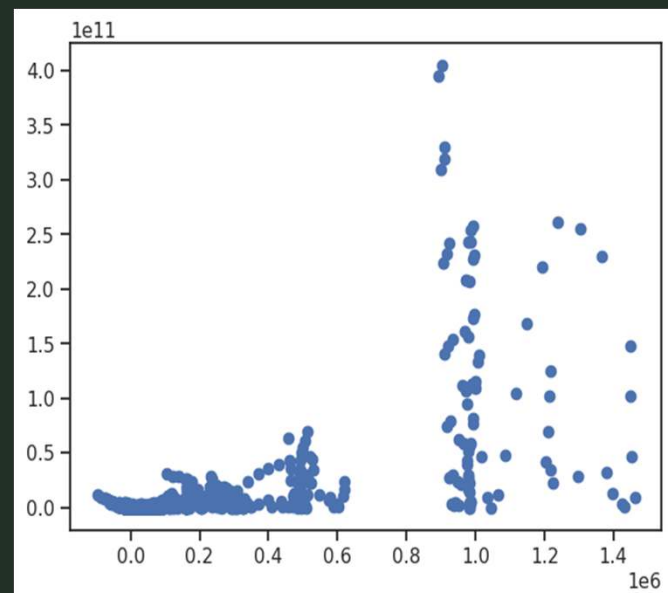
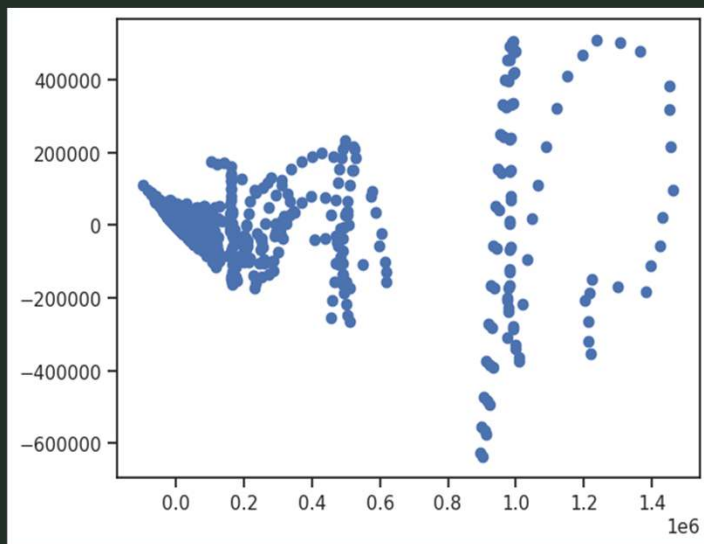
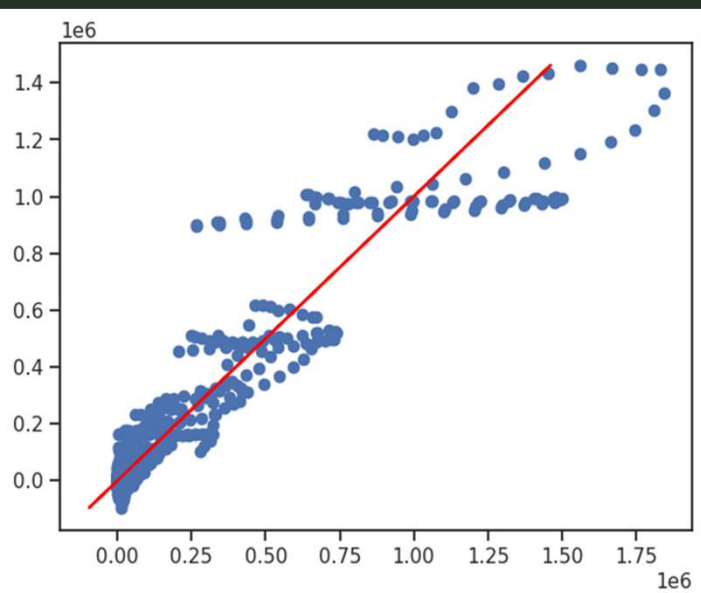
CatBoostRegressor

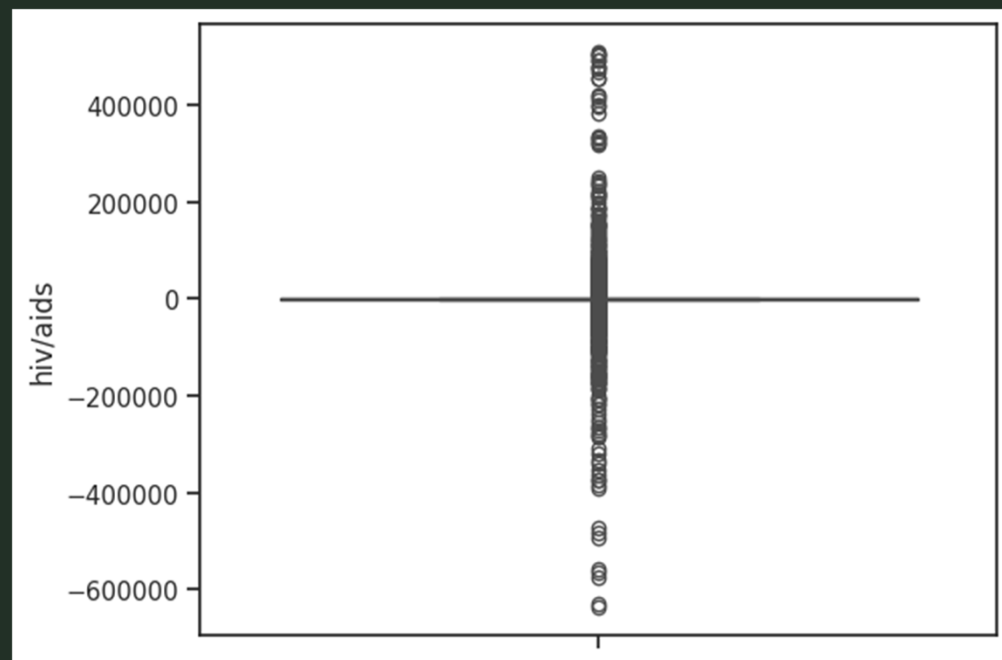
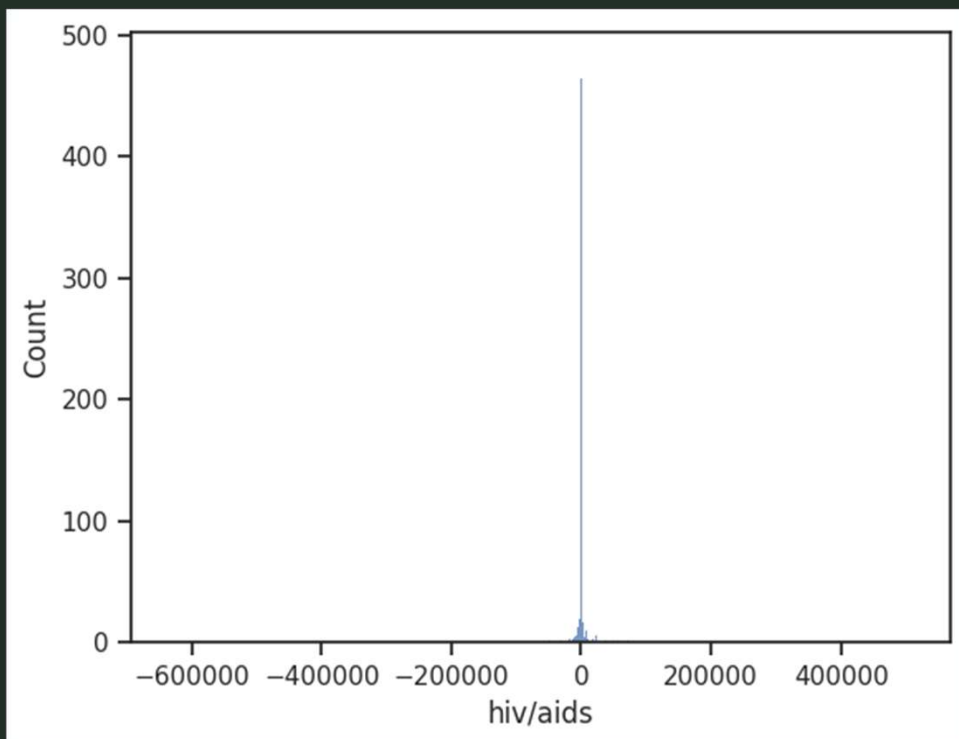
```
➡ MEAN: 151.5640006565424  
MSE: 519542.1577725384  
RMSE: 720.79272871786  
MAPE: 7853421997453755.0
```

Se optimiza modelo con random forest classifier:

```
Best Parameters: OrderedDict([('max_depth', 2), ('n_estimators', 1980)])
```

VIH/AIDS





Se entreno y testeo el modelo obteniendo las siguientes métricas:

XGBRegressor

```
MEAN: 2174.1281206407484  
MSE: 138249573.7669016  
RMSE: 11757.957891015838  
MAPE: 2089391133519532.5
```

LGBMRegressor

```
MEAN: 2823.0900579155405  
MSE: 150909813.63562122  
RMSE: 12284.535548225713  
MAPE: 1.2951091151090203e+17
```

CatBoostRegressor

```
MEAN: 1846.9306301086556  
MSE: 43098090.16512423  
RMSE: 6564.913568747438  
MAPE: 6.120570087948382e+16
```

Se optimiza modelo con random forest classifier:

```
Best Parameters: OrderedDict([('max_depth', 4), ('n_estimators', 97)])
```

Conclusión

Luego de la evaluación se puede concluir:

- En algunas enfermedades hubo aumento de casos de muerte y en otras se mantuvo estable, quedaría evaluar si los avances tecnológicos permitieron un retroceso de las mismas o permitirán un retroceso.
- Para poder evaluar el modelo se utiliza regresión lineal, se usa la optimización y se evalúan las métricas, obteniendo LGBRegressor como la mejor.
- Es posible decir que se puede predecir la evolución de las causas de muertes en Argentina a lo largo de los años con modelos de regresión lineal.