

INCIDENCIA DEL CANCER A NIVEL GLOBAL Y EUROPA

Fernando Moro Cordobés

Introducción

El cáncer constituye una de las principales causas de mortalidad en los seres humanos. Recientemente se ha puesto de manifiesto que el número de personas diagnosticadas con algún tipo de cáncer ha [incrementado a nivel global](#).

Más allá de la búsqueda de nuevos métodos de diagnóstico y terapias, el análisis de los datos de la enfermedad en las distintas zonas del planeta contribuye al estudio relacional de los factores diferenciales en aquellas poblaciones donde existe mayor incidencia. Además, arroja una imagen de aquellos lugares donde es más necesaria la financiación de la atención oncológica y paliativa.

El proyecto [Cancer Incidence in Five continents \(CI5\)](#) es una colaboración entre *International Agency for Research on Cancer* y *International Association of Cancer Registries* cuyo objetivo es recopilar los datos de incidencia del cáncer en distintos países a lo largo del tiempo. En este trabajo se ha utilizado el dataset del proyecto CI5 para realizar una comparativa del número de casos detectados a lo largo de los últimos años en los distintos continentes. Además se han analizado datos a nivel del continente europeo, profundizando en aquellos países donde el número de casos fue más alto.

Para cada región, el número de casos para un tipo de cáncer específico es expresado como el número de casos por cada 100000 habitantes. En el dataset original existen regiones diferentes para un mismo país. Como no conocemos los habitantes de cada región no podemos hacer un promedio ponderado de los casos, así que vamos a trabajar con el número de casos acumulados para reflejar la carga total de la enfermedad en distintas resoluciones espaciales.

Código y Resultados

Importación y preparación del dataset inicial

Out[1]:

	id_code	sex	cancer_code	age	cases	py	year
0	80000299	1	1	1	30	96307.0	1993
1	80000299	1	1	2	18	66677.0	1993
2	80000299	1	1	3	11	59556.0	1993
3	80000299	1	1	4	8	60462.0	1993
4	80000299	1	1	5	30	72770.0	1993
5	80000299	1	1	6	46	64952.0	1993
6	80000299	1	1	7	58	45156.0	1993
7	80000299	1	1	8	37	28283.0	1993
8	80000299	1	1	9	28	18323.0	1993
9	80000299	1	1	10	21	12196.0	1993

```

<class 'pandas.core.frame.DataFrame'>
Index: 21884148 entries, 19 to 23236618
Data columns (total 7 columns):
#   Column      Dtype
---  -
0   cases       int64
1   year        int64
2   name        object
3   region      object
4   cancer_label object
5   edad        object
6   sexo        object
dtypes: int64(2), object(5)
memory usage: 1.3+ GB

```

Out[3]:

	region	year	cantidad_paises
0	Africa	1993	1
1	Africa	1994	1
2	Africa	1995	1
3	Africa	1996	1
4	Africa	1997	1
...
225	Oceania	2013	2
226	Oceania	2014	2
227	Oceania	2015	2
228	Oceania	2016	2
229	Oceania	2017	2

230 rows × 3 columns

Análisis de los datos a nivel de continente

Pregunta 1. Existen diferencias entre los continentes respecto al número de casos diagnosticados de diferentes tipos de cáncer.

Hipótesis: El número de casos de cáncer diagnosticados difiere entre los continentes a lo largo de los años.

Métodos y resultados

1. Cálculo del periodo donde existen datos de al meos un país para los 5 continentes

Con eso comprobamos que el periodo para el que existen datos para todos los continentes es desde 1993 hasta 2017, por lo que trabajaremos con ese periodo.

2. Estudiamos los tipos de cáncer que acumulan mayor número de casos para cada continente en el periodo 1993-2017

C:\Users\ferna\AppData\Local\Temp\ipykernel_12536\2766446274.py:12: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.
Try using `.loc[row_indexer,col_indexer] = value` instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy

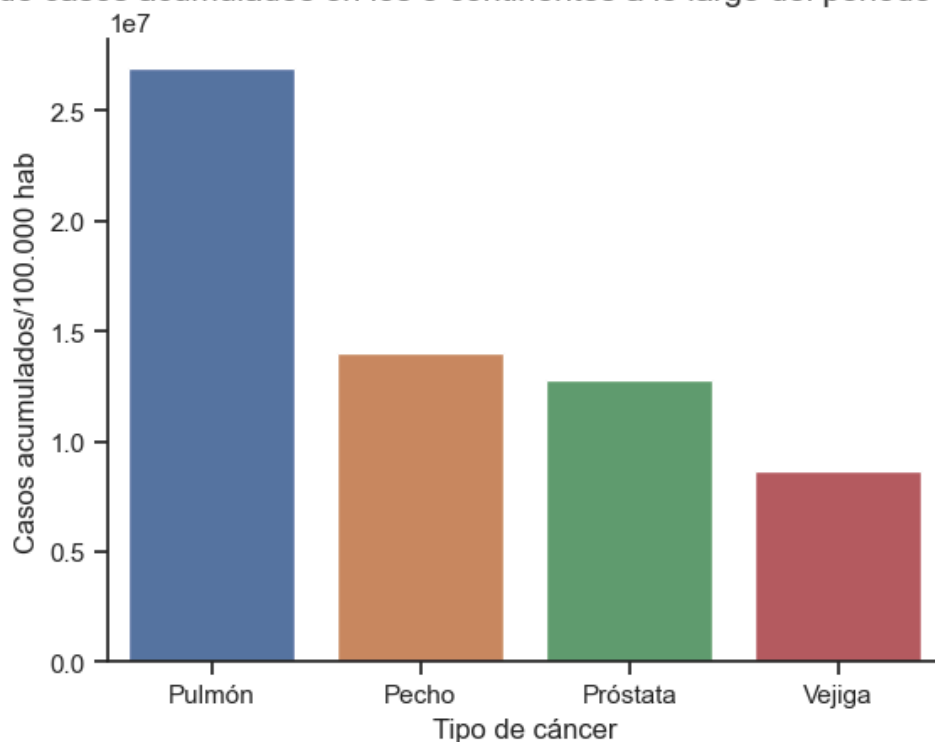
Out[5]:

	cancer_label	cases
0	Pulmón	26926226
1	Pecho	14015528
2	Próstata	12771689
3	Vejiga	8642728
4	Colon	7443348
5	Thyroid	5462162
6	Corpus uteri	5328905
7	Liver	4184982
8	Melanoma of skin	3859782
9	Non-Hodgkin lymphoma	3657172
10	Rectum	3423027
11	Ovary	3118043
12	Stomach	2989148
13	Kidney	2933950
14	Lymphoid leukaemia	2834026
15	Brain, central nervous system	2829600
16	Pancreas	2688141
17	Oesophagus	2561928
18	Myeloid leukaemia	2463406
19	Cervix uteri	2251449
20	Multiple myeloma	1309966
21	Testis	1122383
22	Hodgkin lymphoma	953374
23	Larynx	852655
24	Gallbladder etc.	823727
25	Connective and soft tissue	674273
26	Anus	667582
27	Tongue	654046
28	Mouth	522668
29	Bone	476316
30	Small intestine	423301
31	Eye	367721
32	Tonsil	367282

	cancer_label	cases
33	Vulva	290960
34	Mesothelioma	266220
35	Immunoproliferative diseases	265697
36	Salivary glands	253521
37	Hypopharynx	196163
38	Renal pelvis	192181
39	Lip	176664
40	Nose, sinuses etc.	157734
41	Leukaemia, cell unspecified	147262
42	Nasopharynx	145548
43	Ureter	128003
44	Other oropharynx	118854
45	Other female genital organs	106384
46	Kaposi sarcoma	106373
47	Uterus unspecified	99209
48	Penis	97859
49	Vagina	80489
50	Adrenal gland	69708
51	Other urinary organs	67038
52	Pharynx unspecified	65903
53	Burkitt lymphoma	61561
54	Meninges	42134
55	Other endocrine	35368
56	Other male genital organs	24755
57	Trachea	15761
58	Placenta	10513

Out[6]: Text(0, 0.5, 'Casos acumulados/100.000 hab')

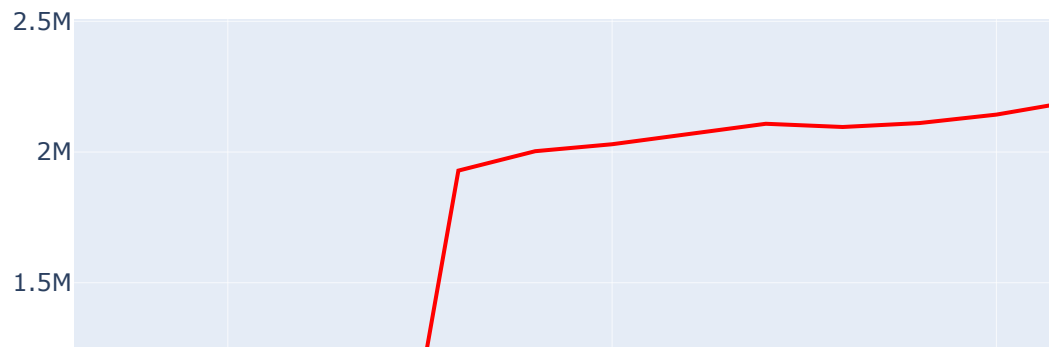
Nº de casos acumulados en los 5 continentes a lo largo del periodo 1993-2017



Comprobamos que a nivel de los 5 continentes, los tipos de cáncer que mayor número de casos acumulan son el cáncer de pulmón, pecho, próstata y vejiga.

3. Estudiamos como ha evolucionado el número de casos de los 4 tipos de cánceres con mayor número en los 5 continentes durante el periodo 1993-2017 y cuántos países dentro de cada continente aportaban información.

Número de casos acumulados por continente y año



Se observa que el continente que tiene un mayor número de países que contribuyen con información es Europa (21).

Confirmamos la hipótesis de que existen diferencias entre el número de casos de cáncer diagnosticados entre los diferentes continentes. Estas diferencias se pueden deber a múltiples factores, como genéticas, estilo de vida, accesibilidad a medios sanitarios, organización en la toma de datos, avances en la tecnología de diagnosis..etc

Al observar que entre el año 1997 y 1998 el número de casos diagnosticados con algún tipo de cáncer en América experimenta un dramático ascenso, buscamos en los datos a que puee ser debido. Hipotetizamos que puede deberse a que países pertenecientes a dicho continente empezaron a aportar información.

```
Out[10]: array(['Chile', 'Colombia', 'Costa Rica', 'Ecuador', 'Canada',
               'United States of America'], dtype=object)
```

Con esto comprobamos que en el año 1998 se suman Chile al conteo de casos. Confirmando que esto puede ser la fuente de incremento repentino en el número de casos diagnosticados.

```
Out[11]: array([2003, 2004, 2005, 2006, 2007, 2008, 2009, 2010, 2011, 2012, 2013,
                2014, 2015, 2016, 2017])
```

Existe una problemática importante. La incorporación de países en diferentes años a lo largo de la serie temporal constituye otra fuente de variabilidad que hace que la comparación de los datos a determinado nivel no sea estadísticamente correcta.

Análisis de los datos a nivel de Europa

Esta gran variabilidad hace que la comparación entre continentes presente un alto grado de error.

A partir de aquí vamos a estudiar las potenciales diferencias en los casos de cáncer dentro del continente europeo a lo largo del periodo entre 1993 y 2017. Indagaremos en cuáles fueron los tipos de cáncer más comunmente diagnosticados a lo largo de dicho periodo, si existen diferencias entre la edad de las personas diagnosticadas y si el sexo es algo que influya en dicho diagnóstico.

1. Analizamos cuáles son los tipos de cánceres más comunes en Europa en el periodo estudiado


```

Out[12]: cancer_label
Pulmón 4840980
Pecho 2605993
Próstata 2222048
Vejiga 1762490
Colon 1548686
Corpus uteri 955935
Rectum 880271
Ovary 727229
Oesophagus 704568
Melanoma of skin 667215
Stomach 658524
Non-Hodgkin lymphoma 589652
Brain, central nervous system 586456
Kidney 544867
Pancreas 542072
Liver 536386
Lymphoid leukaemia 511115
Cervix uteri 488689
Thyroid 450428
Myeloid leukaemia 424900
Testis 269031
Multiple myeloma 241610
Larynx 187770
Hodgkin lymphoma 184440
Gallbladder etc. 158438
Mouth 116578
Connective and soft tissue 113636
Anus 108678
Tongue 104645
Mesothelioma 90094
Bone 83664
Eye 80425
Vulva 70618
Tonsil 62301
Small intestine 61954
Hypopharynx 57094
Lip 43102
Salivary glands 39486
Renal pelvis 36329
Immunoproliferative diseases 35335
Other oropharynx 34008
Nose, sinuses etc. 32473
Penis 29315
Leukaemia, cell unspecified 26206
Ureter 22799
Other female genital organs 20353
Nasopharynx 18534
Uterus unspecified 16730
Vagina 14946
Adrenal gland 13862
Pharynx unspecified 12369
Other urinary organs 12033
Kaposi sarcoma 9979
Meninges 9475
Burkitt lymphoma 8664
Other endocrine 6105
Trachea 4867
Other male genital organs 4404

```

Placenta

1275

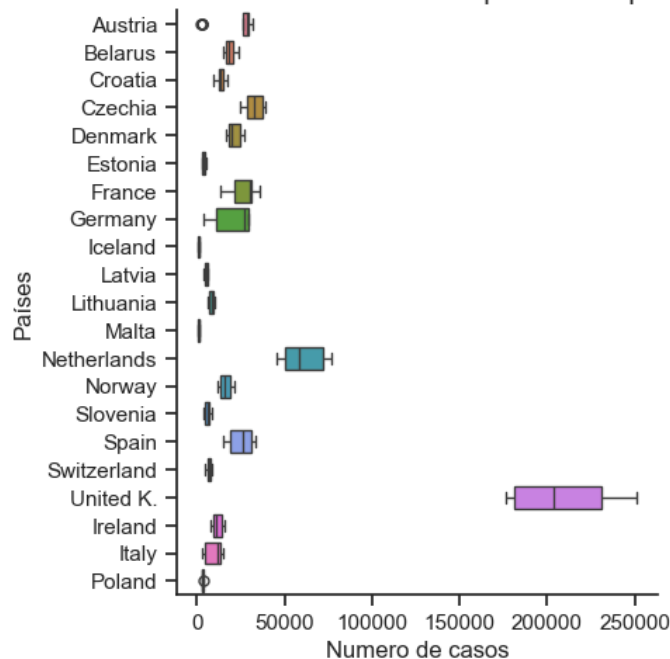
Name: cases, dtype: int64

Los cánceres más comunes con los que trabajaremos a partir de ahora son: Pulmón, Pecho, Próstata, Vejiga, Colon. Para visualizar como se distribuyen los conteos de casos a lo largo del periodo realizamos una representación tipo boxplot de los países del continente Europeo.

2. Analizamos como se distribuyen el número de cánceres más comunes a lo largo del periodo en los diferentes países, con la finalidad de ver si existen diferencias entre países Europeos

<Figure size 1500x3000 with 0 Axes>

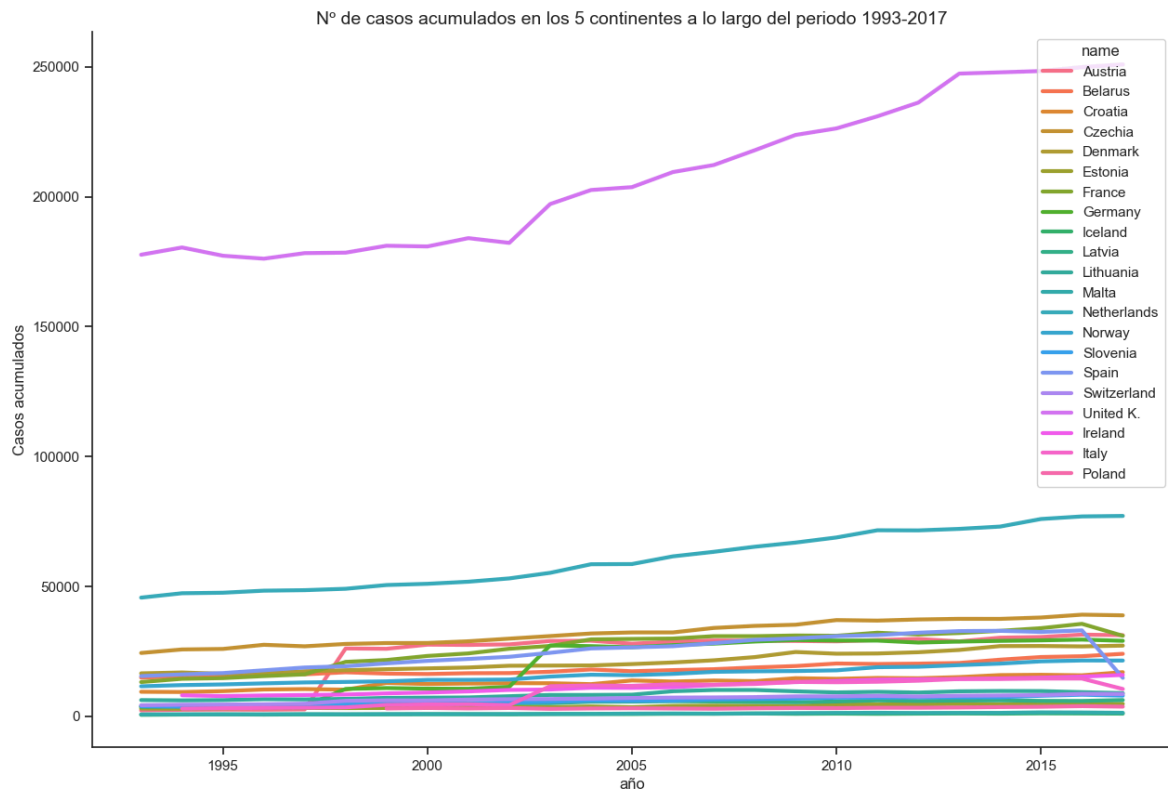
Distribución de casos de cánceres más comunes en países Europeos en el periodo 1993-2017



El gráfico de caja muestra como se distribuyen el numero de casos a lo largo del periodo de enlos diferentes países. Se observa en Netherlands (Holanda) y el reino unido (United K.) se diagnostican mayor número de canceres.

Vamos a representar gráficamente como han evolucionado el número de diagnósticos de los cánceres más comunes en los países europeos a lo largo de los años en el periodo 1993-2017.

3. Estudiamos como han evolucionado el número de cánceres diagnosticados en cada uno de los países de Europa



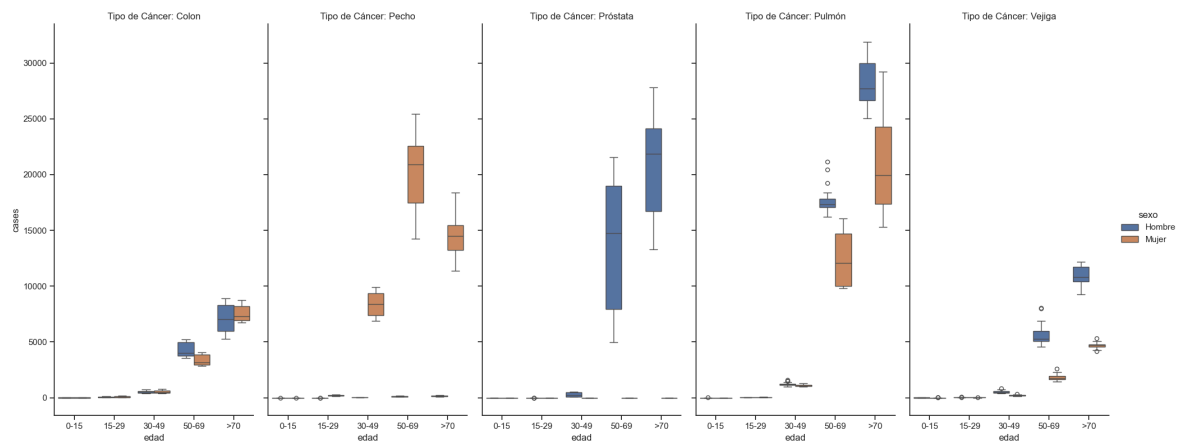
El análisis gráfico muestra como el Reino Unido y Holanda son los países donde más casos de los 5 tipos de cánceres se han diagnosticado en el periodo entre 1993 y 2017. Además se observa una tendencia en el aumento de casos en el periodo señalado en varios países.

Análisis de los datos de Reino Unido y Holanda

Dado el alto número de casos en Reino Unido y Holanda, vamos a estudiar como se distribuyen los promedios de casos por año para cada uno de los 5 tipos de cánceres más comunes. Primero lo analizaremos por edad y sexo, y después analizaremos como ha evolucionado a lo largo de los años del periodo 1993-2017.

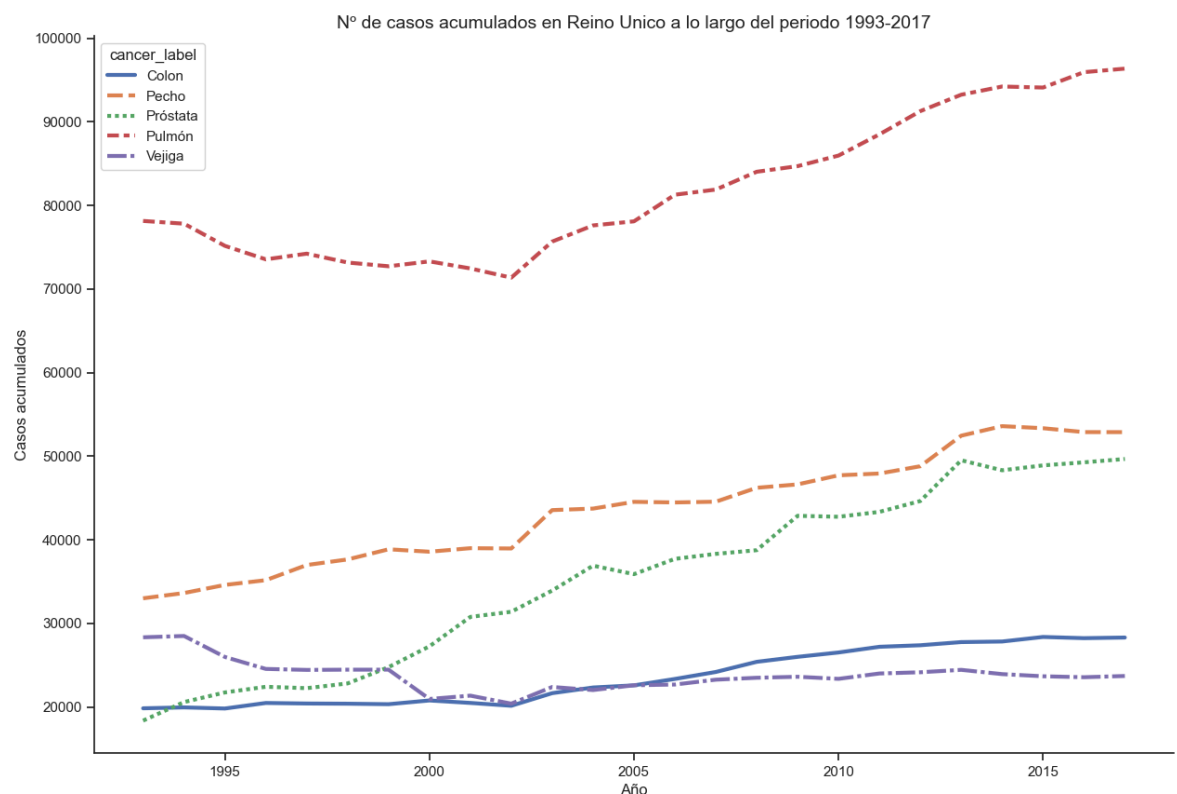
Análisis de los datos en Reino Unido

1. Analizamos el número de casos acumulados a lo largo del periodo 1993-2017 en el Reino Unido para cada uno de los cánceres más comunes en los diferentes rangos etarios.



Podemos observar en los gráficos que en todos los tipos de cancer el número de casos aumenta con la edad. Además es de destacable que el número de cánceres de pecho diagnosticados en hombres es prácticamente 0. Como es lógico, no existen casos de cáncer de próstata en mujeres. La mayor variabilidad en el número de casos a lo largo del periodo ocurrió en los diagnosticados de cáncer de próstata.

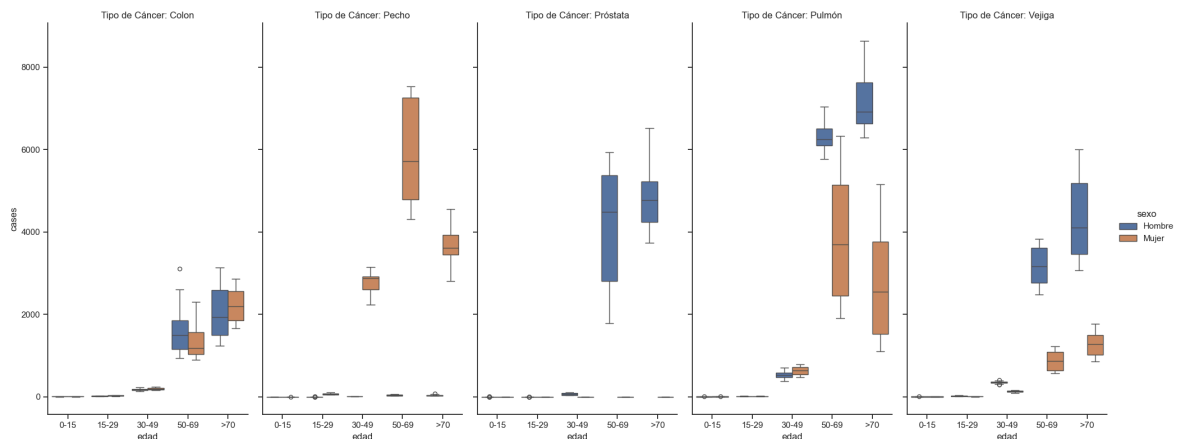
2. Estudiamos la evolución de casos diagnosticados de cada tipo de cáncer en Reino Unido a lo largo del periodo 1993-2017



Podemos observar que el número de casos aumenta a lo largo del periodo en todos los tipos de cáncer excepto en el de vejiga.

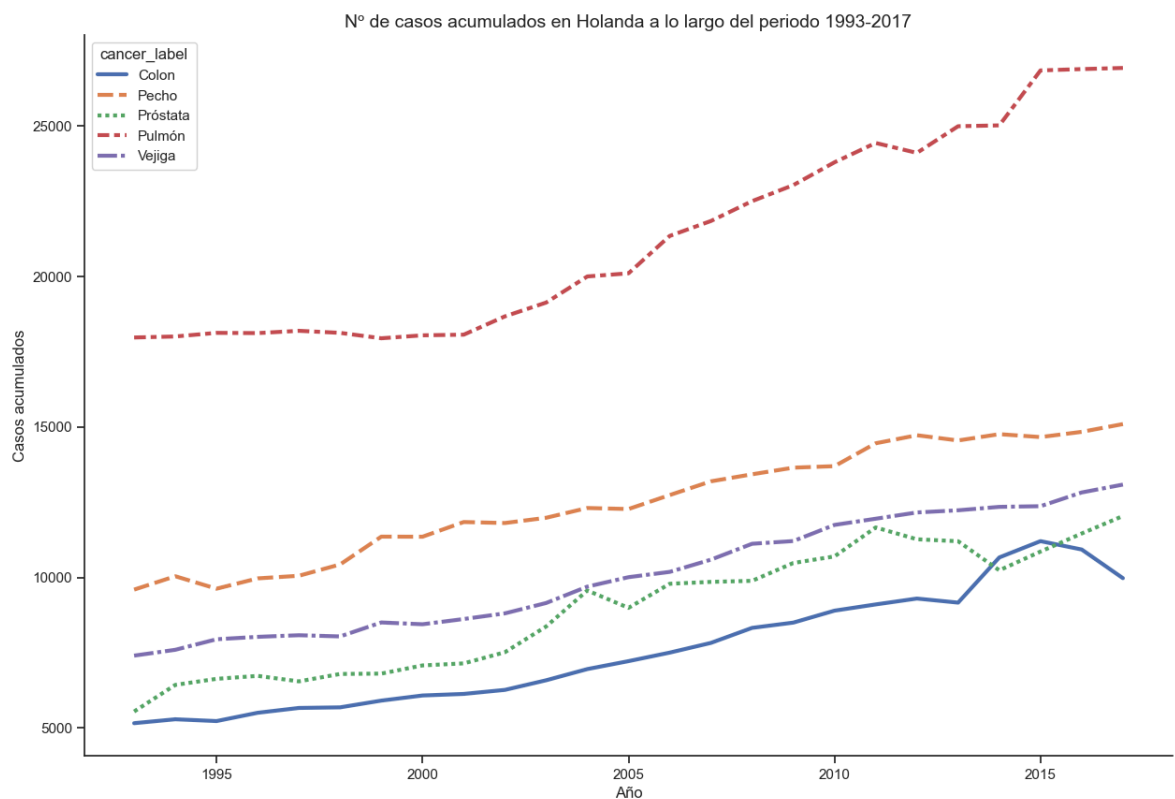
Análisis de los datos en Holanda

1. Analizamos el número de casos acumulados a lo largo del periodo 1993-2017 en el Reino Unido para cada uno de los cánceres más comunes en los diferentes rangos etarios.



Al igual que en UK vemos como el número de casos diagnosticados de los diferentes tipos de cáncer se acumulan en los rangos de edad a partir de 30 años. También se observa la disparidad de cifras entre hombres y mujeres en los casos de cánceres de hombres y mujeres.

2. Estudiamos la evolución de casos diagnosticados de cada tipo de cáncer en Reino Unido a lo largo del periodo 1993-2017



En el gráfico se observa como en Holanda el número de casos de todos los cánceres analizados han aumentado a lo largo del periodo 1993-2017.

Conclusión

- Existen diferencias en la incidencia del cáncer entre continentes en el periodo 1993-2017. Aun así la información está muy sesgada debido a que el número de países que contribuyen con información en cada continente presenta una desigualdad bastante pronunciada.
- El número de casos acumulados a nivel global es mayor en el caso de los cánceres de pulmón, pecho, próstata y vejiga. Además parece existir cierta tendencia en alza en el número de casos acumulados a lo largo del periodo estudiado.
- Los tumores más comunes en Europa son de pulmón, pecho, próstata, vejiga y colon.
- Los países con mayor número de casos acumulados de los tumores más comunes son el Reino Unido y Holanda.
- El número de casos de los tumores más comunes en Reino Unido y Holanda aparecen presentan las cifras más altas en rangos de edad superiores a los 30 años.
- El número de casos diagnosticados de los tipos de cánceres más comunes muestran una tendencia en alza en Europa en el periodo de años estudiado.