

Análisis de Datos - Tarea 4

María José Fonseca Vázquez

21 de Febrero de 2020

Abstract

Se realizaron distintos gráficos generados a partir de una base de datos que contiene el número total y la tasa de prisioneros por país para condensar toda la información en gráficos entendibles y de fácil interpretación.

Tasa de Prisioneros

Teníamos tres bases de datos a elegir. En mi caso seleccioné el llamado *prison rates*, que contenía información sobre los prisioneros que, además de mencionar el país, contenía la región -continente- y subregión. Toda la información estaba separada por años, comenzando en 2003 y terminando en 2017.

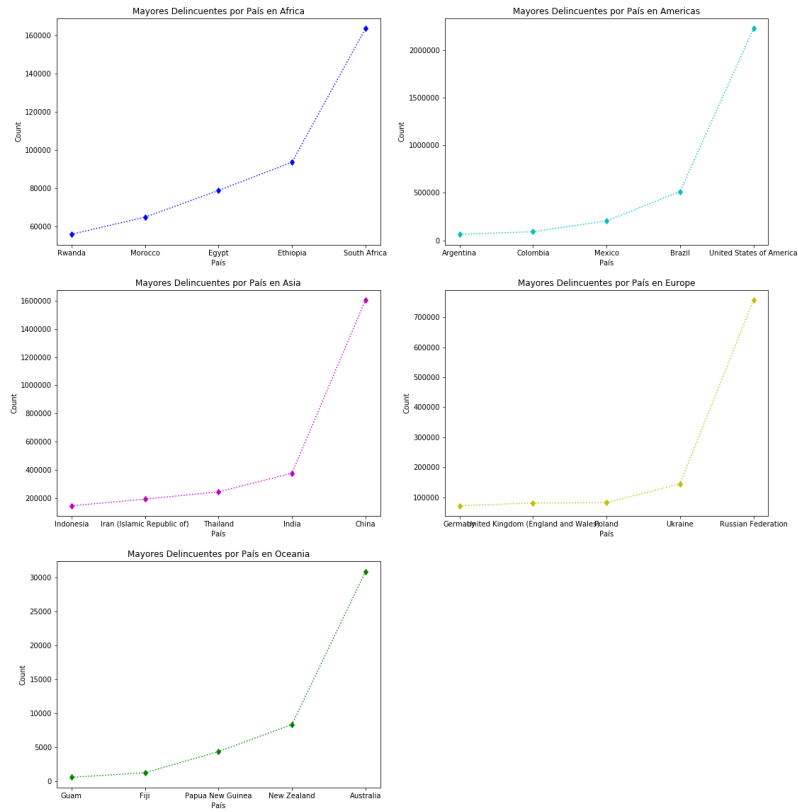
El problema de esta base de datos fue que había datos inexistentes, ya que para algunos países no aparecía información de cada uno de los años. Sin embargo, era sencillo trabajar con ella a pesar de todo, ya que si no había información, no aparecían columnas con campos vacíos, sino que simplemente no aparecía la fila con ninguna casilla, así que se podía realizar una *máscara* de manera muy sencilla y trabajar con la base de datos tal cual.

¿Cuál es el país -por continente- con mayor número promedio de prisioneros?

Para este caso, se realizó para cada continente una base de datos que tuviera sólo la información del mismo y se agruparon los datos por país, de tal manera que se tuviera para cada país el promedio de prisioneros a lo largo de los años. Como eran muchos países por mes y lo que necesitamos saber es el país con mayor número de personas detenidas, sólo grafiqué para cada caso los cinco países con mayor valor. Se observa una distribución semejante en cada caso, a pesar de que se trate de continentes diferentes, pues siempre hay un país que predomina por mucho respecto a los demás. En la Figura 1 se muestran las gráficas obtenidas para esta pregunta.

Entonces, podemos ver fácilmente que la respuesta a la pregunta es:

Figure 1: Prisioneros Promedio por Continente

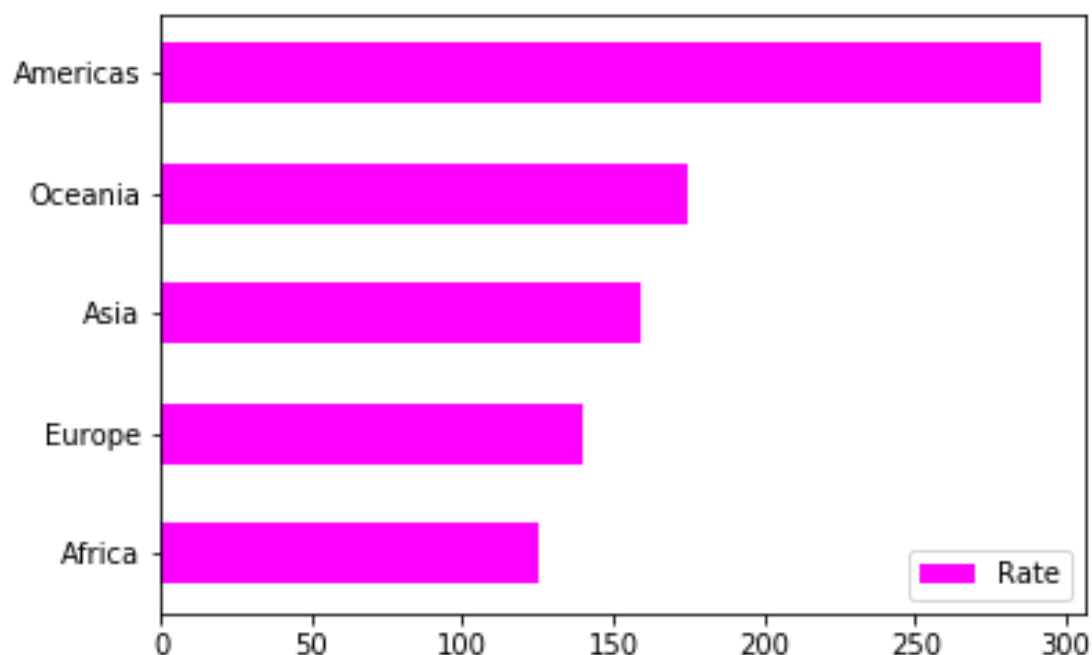


- Africa → South Africa
- América → USA
- Asia → China
- Europa → Federación Rusa
- Oceanía → Australia

¿Cuál es el continente con mayor tasa promedio de prisioneros?

Para este caso, agrupé por continente la base de datos original, y al agruparla calculé el promedio de todas las celdas, obteniendo así la tasa promedio de delincuentes por continente. Se puede apreciar fácilmente en la Figura 2 que el continente que ocupa el primer lugar es Estados Unidos.

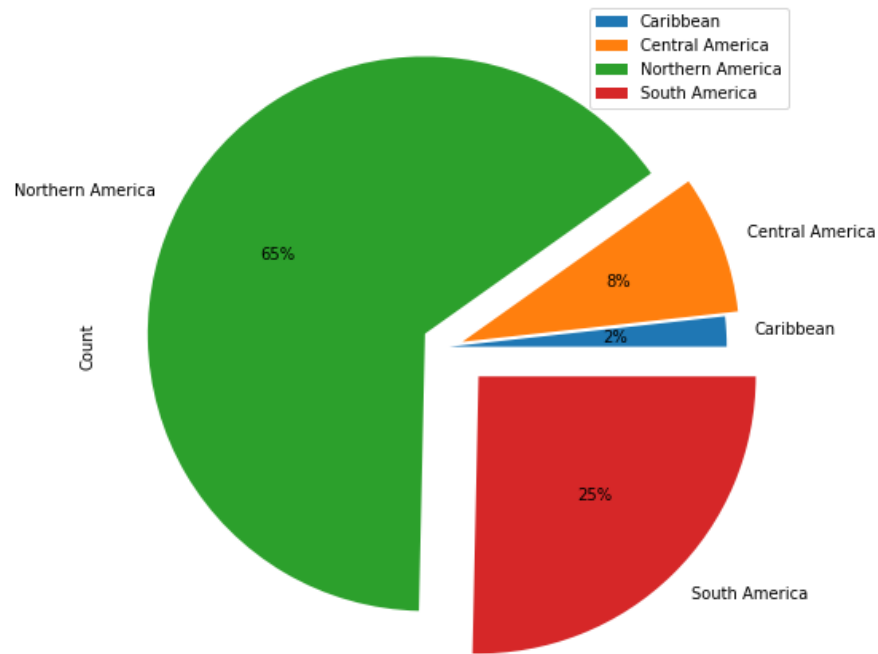
Figure 2: Tasa Promedio por Continente



¿Qué subregión de dicho continente tiene la mayor cantidad de prisioneros?

Para este caso, filtré la base de datos para el continente americano y, como lo que importaba era únicamente ver a qué región le correspondía una mayor cantidad de prisioneros, hice una gráfica de pastel, como se muestra en la Figura 3, y se observa que la cantidad de prisioneros se encuentra mayormente concentrada en América del Norte.

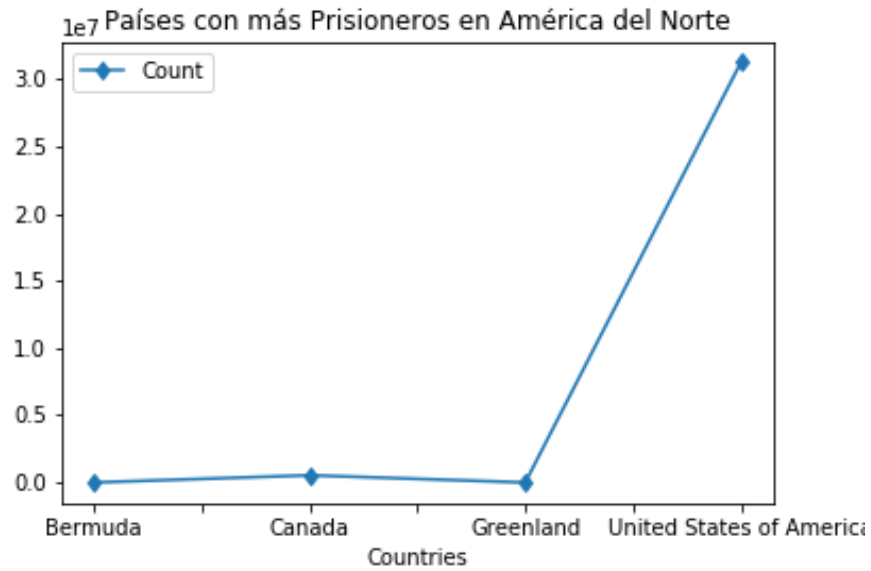
Figure 3: Subregiones del Continente Americano



¿Cuál es el país en esa región con la mayor cantidad de prisioneros?

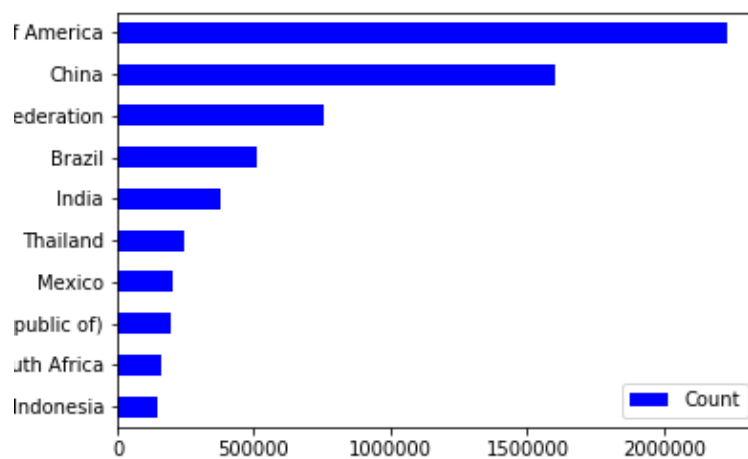
Al igual que en el caso anterior, filtré los datos por continente americano y región América del Norte, sumando toda la cuenta de prisioneros a lo largo del tiempo. El resultado era de esperarse, ya que si vemos la Figura 1 sabemos que el país que debería dominar en la gráfica es Estados Unidos de América, como se observa en la Figura 4.

Figure 4: Prisioneros en América del Norte



¿Cuáles son los 5 países alrededor del mundo con mayor número de prisioneros?

Figure 5: Países con más Prisioneros a Nivel Mundial

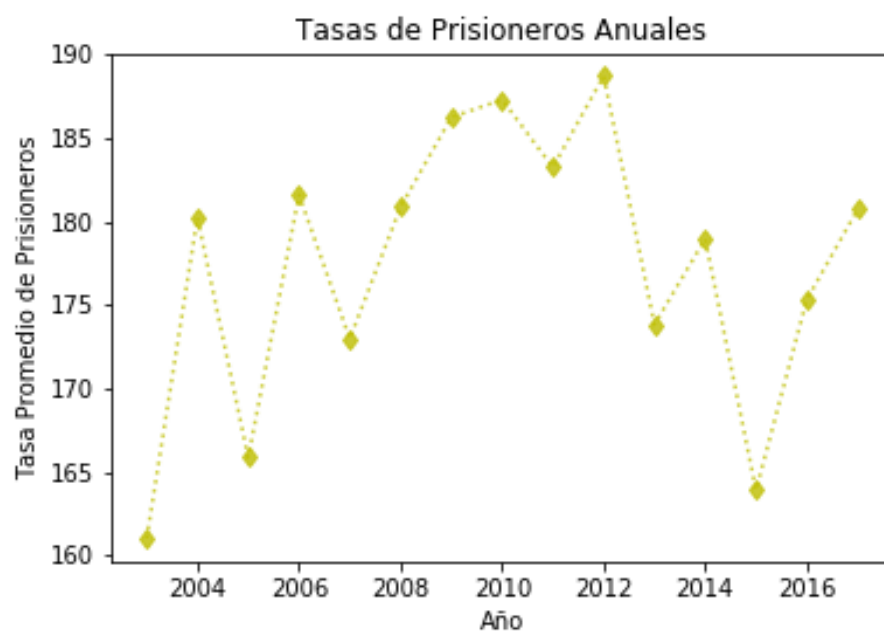


Para este caso, agrupé los datos por país con el promedio a lo largo de los años y grafiqué sólo los 10 con mayor número. De acuerdo a la Figura 5, los países que ocupan los 5 primeros puestos son USA, China, la Federación Rusa, Brasil e India.

¿Cómo ha aumentado la tasa de prisioneros a nivel mundial con el paso de los años?

Para este caso, filtré los datos por año y calculé la tasa promedio mundial. Se observan aumentos y decrementos, es decir, no se sigue un patrón ni se tiene una periodicidad, como se observa en la Figura 6.

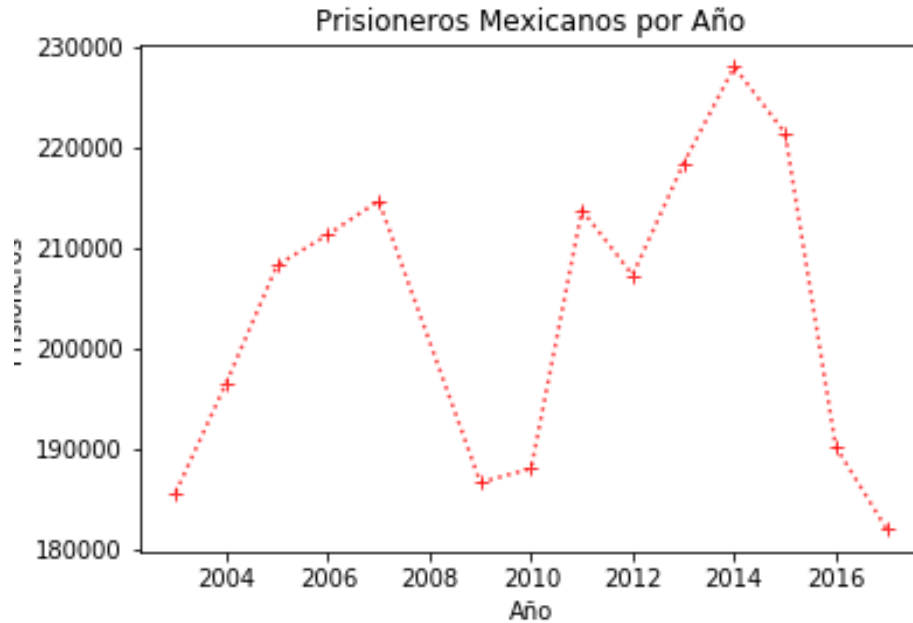
Figure 6: Tasa Anual Mundial con el Paso de los Años



¿En qué año tuvo México la mayor cantidad de prisioneros?

Se filtró la base de datos para que se tuviera sólo la información para México y se graficó la cantidad de prisioneros respecto al año. Los resultados se muestran en la Figura 7

Figure 7: Tasa Anual Mundial con el Paso de los Años



¿Cuál ha sido el país con mayor cantidad de prisioneros para cada uno de los años?

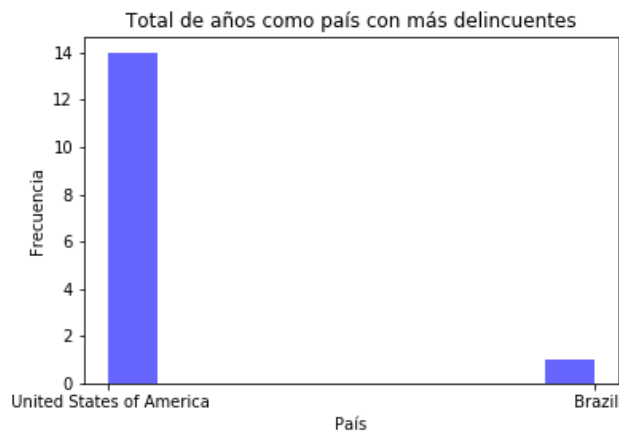
Para esto filtré los datos por año y guardé en un arreglo el nombre del país con mayor cantidad de prisioneros. Para observar el país por año, me pareció más conveniente hacer una tabla, y al final hacer un histograma por país. La información se muestra en la Tabla 1 y en la Figura 8.

Cabe señalar que Estados Unidos no fue el país con más prisioneros sólo para el 2017, pero es probablemente debido a que no hubo datos para ese año.

Table 1: Países con más prisioneros por año.

Año	País con más prisioneros
2003	United States of America
2004	United States of America
2005	United States of America
2006	United States of America
2007	United States of America
2008	United States of America
2009	United States of America
2010	United States of America
2011	United States of America
2012	United States of America
2013	United States of America
2014	United States of America
2015	United States of America
2016	United States of America
2017	Brazil

Figure 8: Frecuencia de países con mayor número de prisioneros



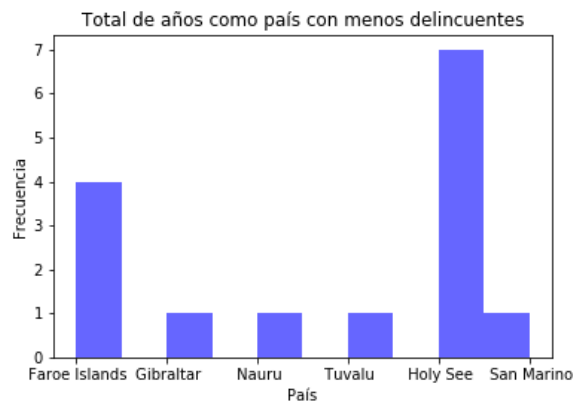
¿Cuál ha sido el país con menor cantidad de prisioneros para cada uno de los años?

Se realizó exactamente lo mismo que en el caso anterior, pero esta vez para menos prisioneros. Holy See fue el país que apareció más veces, con una frecuencia de 7. Los resultados obtenidos se observan en la Tabla 2 y en la Figura 9.

Table 2: Países con más prisioneros por año.

Año	País con menos prisioneros
2003	Faroe Islands
2004	Gibraltar
2005	Nauru
2006	Faroe Islands
2007	Tuvalu
2008	Faroe Islands
2009	Holy See
2010	Holy See
2011	Holy See
2012	Holy See
2013	Holy See
2014	Holy See
2015	Holy See
2016	Faroe Islands
2017	San Marino

Figure 9: Frecuencia de países con mayor número de prisioneros



Conclusiones

Pandas es muy útil, ya que es muy fácil realizar filtros a cualquier base de datos, además de que es sencillo agrupar, sacar promedios, máximos, entre otros. Se puede realizar cualquier gráfico, aunque necesitamos saber distinguir cuál es el mejor, ya que en algunos es más fácil apreciar cierta información.