

# PRA 2: Creación de la visualización y entrega del proyecto

## Medidas de control sobre el tabaquismo aplicadas en diferentes países y regiones

Autor: Adonis González Godoy

Junio 2021

### Contents

|   |   |
|---|---|
| 1. Título . . . . .                               | 2 |
| 1.1 URL . . . . .                                 | 2 |
| 1.2 Summary . . . . .                             | 2 |
| 1.3 Los datos . . . . .                           | 2 |
| 1.4 Descripción . . . . .                         | 2 |
| 1.5 Las cuestiones de la visualización . . . . .  | 3 |
| 2. Exploración de los datos . . . . .             | 3 |
| 3. Visualisations . . . . .                       | 4 |
| 4. Conclusiones . . . . .                         | 4 |
| 4.1 Conclusiones de las visualizaciones . . . . . | 4 |
| 5. Descripción técnica del proyecto . . . . .     | 5 |
| 6. Files . . . . .                                | 5 |
| Author . . . . .                                  | 5 |

## 1. Título

### Medidas de control sobre el tabaquismo aplicadas en diferentes países y regiones

#### 1.1 URL

- Tableau public

#### 1.2 Summary

En este proyecto de visualización se intentará dar respuestas a una serie de preguntas relacionadas con el consumo de tabaco a nivel mundial a partir de la información recogida por la Organización Mundial de la Salud (WHO). Con el fin de intentar mejorar el bienestar de las personas y reducir los riesgos de la salud asociados al consumo de tabaco, contribuyendo así a mejorar la salud de la población.

#### 1.3 Los datos

El dataset ha sido seleccionado de la página oficial de la Organización Mundial de la Salud. La información se refiere al Tabaquismo en el año 2016 y hace un seguimiento de la situación en diferentes países y regiones que han aplicado políticas de control del tabaco que van desde advertencias gráficas y prohibiciones de publicidades hasta zonas de no fumadores. Los datos se han recopilado para intentar mejorar el bienestar de las personas y reducir los riesgos de la salud asociados al consumo de tabaco, contribuyendo así a mejorar la salud de la población.

Links del dataset:

- Dataset (.csv)
- Official dataset source

#### 1.4 Descripción

El tabaquismo es un trastorno crónico en que la persona presenta una adicción física como psicológica al consumo de tabaco. El consumo de tabaco es uno de los principales factores de riesgo carcinógeno y otras enfermedades pulmonares. Además de ser altamente adictivo, la nicotina es la droga responsable en primer lugar de la adicción de una persona, esta adicción es semejante a la adicción producida por el uso de drogas como la heroína y la cocaína. El tabaco también puede ser mortal para los no fumadores. La exposición al humo de segunda mano también se ha visto implicada en resultados adversos para la salud. Fumar durante el embarazo puede provocar varias afecciones de por vida en los bebés. En este proyecto de visualización se intentará dar respuestas a una serie de preguntas relacionadas con el consumo de tabaco a nivel mundial a partir de la información recogida por la Organización Mundial de la Salud (WHO). Con el fin de intentar mejorar el bienestar de las personas y reducir los riesgos de la salud asociados al consumo de tabaco, contribuyendo así a mejorar la salud de la población.

## 1.5 Las cuestiones de la visualización

Se intentará responder a las siguientes preguntas:

- ¿Qué países/regiones tienen el precio más alto de tabaco?
- ¿Qué países/regiones tiene los impuestos más altos sobre el tabaco?
- ¿Qué países/regiones tienen más restringidas las áreas de fumadores?
- ¿Qué países/regiones tiene las políticas de control de tabaco más elevadas?
- ¿Qué países/regiones tiene prohibiciones de publicidad en los medios de comunicación sobre el tabaco?

En este proyecto se dará respuesta a las preguntas citadas con diferentes visualizaciones en un dashboard. Se utilizará la herramienta Tableau.

## 2. Exploración de los datos

El primer contacto con el dataset ha sido realizado usando Jupyter Notebook en un conda environment que ejecuta la versión 3.6 de Python. Algunas de las librerías que se han usado son: Padas, Numpy, Matplotlib.

De la exploración de los datos se ha obtenido:

- El dataset contiene 173 variables por los 194 países entre las diferentes regiones registrados. La dimensión del dataset es de **173x194**.
- El dataset es variado y tiene variables combinadas entre sus diferentes tipos, tiene variables categóricas, numéricas, cadena de caracteres y variables de punto flotante.
- El dataset contiene datos nulos o “-” de los países que no se han recogido datos.
- Se han realizado transformaciones de variables detectadas como string a numeric.

La exploración de datos se encuentra disponible en:

- EDA - HTML
- EDA - Jupyter Notebook
- Download EDA - Jupyter Notebook

### 3. Visualisations

- Tableau public

Respuestas:

En el primer gráfico podemos observar el precio del tabaco en los diferentes países, en concreto el precio de 20 cigarrillos de la marca más vendida en dólares en tipo de cambio oficial, se aprecia a mayor intensidad de color el precio es más elevado, a partir de aquí damos respuesta sobre los países y regiones con los precios más elevados, Western pacific parece tener los precios más elevados del tabaco, en concreto se puede observar Australia y Nueva Zelanda con el precio por encima de los 15, *porotroladolaregióndeÁfricapareceserlaregiónconlospreciosmásbajosdeltabaco, alrededor de 1,5* entre sus países y siendo Somalia el país con el precio más bajo de esta región a solo 46 centavos de dólar los 20 cigarrillos.

En la región de Europa Noruega es el país con el precio más alto por encima de los 12\$ mientras que República Checa tiene los precios más bajos 3,45, *Enespañaelprecioseencuentraporencimadelos5*.

En América, los precios más altos se localizan en el norte, en concreto en Estados Unidos y Canadá por otro lado los más bajos en Colombia a tan solo 0,88\$ (De Venezuela no hay datos recogidos sobre el precio del tabaco).

En el gráfico de burbujas agrupadas tenemos respuesta a las restricciones que tienen respecto al tabaco en la televisión, la radio y los medios impresos nacionales. Como se puede observar clamaramente la región de Europa es la que aplica más prohibición de todas las formas de publicidad directa e indirecta. También se puede ver claramente que una de las regiones con menos prohibición es la de South-East Asia o tiene Ausencia, en otras palabras, en los países del South-East Asiático aún hay publicidad de tabaco ya sea en radio tv o periódicos nacionales.

En el gráfico de tablas resaltadas se da respuesta por cada región, el tipo de restricción que tienen con respecto a los lugares públicos libres de humo. Por ejemplo se puede observar que en la región de América (AMR) tienen una restricción mayor de todos los lugares públicos completamente libres de humo (o al menos el 90% de la población cubierta por una legislación completa de ambientes libres de humo), mientras que África es la región con menos restricciones con respecto a fumar en lugares públicos.

En el gráfico de barras horizontales se muestra el impuesto en porcentaje % del precio de la marca más vendida, Europa es la región donde el impuesto sobre el tabaco es más alto con un 74,80% de media. Por otro lado el impuesto más bajo es en la región de África con un 35,27%.

### 4. Conclusiones

#### 4.1 Conclusiones de las visualizaciones

- África es la región con el precio más bajo, en concreto Somalia con solo 0,46\$ los 20 cigarrillos de la marca más vendida.
- Australia y Nueva Zelanda tienen el precio más alto de tabaco superando los 15\$
- Europa no solo tiene los impuestos más altos si no también es la región con más importantes restricciones, desde la publicidad prohibida en los medios de comunicación hasta restricciones de fumadores en espacios públicos.
- África es la región con menos restricciones en espacios públicos y además con menos impuestos.

## 5. Descripción técnica del proyecto

En este documento se presenta desde el primer contacto con el conjunto de datos hasta las conclusiones de las visualizaciones. El exploratorio de los datos ha sido realizado usando Jupyter Notebook en un conda enviroment que ejecuta la versión 3.6 de Python. Las librería que se han usado se encuentran en el informe del Jupyter Notebook, entre las principales se tiene que destacar, Padas, Numpy, Matplotlib, Seaborn.

Lo primero que se ha realizado es la carga del dataset para realizar una primera inspección, dimensión, valores nulos, tipo de datos. Cuando se ha requerido se ha transformado para facilitar la manipulación y/o creación de visualizaciones.

La herramienta que se ha utilizado para poder hacer las representaciones gráficas del dashboard es Tableau, esta herramienta permite expresar datos visualmente al transformar acciones de arrastrar y soltar en consultas de dados, en una interfaz intuitiva. Puede permitir incluso aprendizaje automático, estadísticos, lenguaje natural y preparación de los datos inteligentes para que sean más eficaces y permitan aumentar la creatividad de las personas que lo utilizan para hacer análisis.

También se han realizado distintas pruebas con Tableau para poder asegurarse del diseño adecuado que se pretende seleccionar, para que el usuario final puede interactuar. En general, se ha tenido en cuenta que la visualización de todos los gráficos sean capaces de entenderse sin poner texto que los acompañen, pero utilizando herramientas elementales que mejora mucho la legibilidad de las visualizaciones como son las etiquetas y las leyendas.

Por último, se ha exportado el dashboard. Se encuentra disponible en el siguiente link:

- Tableau public

## 6. Files

- Tableau .twb
- Tableau public
- Dataset (.csv)
- EDA - HTML

## Author

Adonis González Godoy