

Análisis sistemático de las tasas de suicidio globales contemporáneas

ASESOR: Dr. Miguel Felix Mata Rivera, Instituto Politécnico Nacional

Joshua Haziél Hernández Castillo, Instituto Politécnico Nacional, joshua.haziel.hc@gmail.com

PLANTEAMIENTO DEL PROBLEMA

El suicidio es un problema de salud social al cual no se le da importancia, de la misma manera la difusión es escasa dando como resultado el no brindar a la sociedad cifras oficiales que permitan analizar dicha causa de mortalidad. Datos proporcionados por la OMS mencionan que la carga de morbilidad producida por dicho problema asciende al 1.4% de la cifras totales a nivel mundial, en otros términos existen aproximadamente 1.5 millones de muertos por esta causa. Un análisis previo del tema permite afirmar que existen factores específicos que influyen en la personalidad de un individuo cuya determinación es el suicidio, es por ello que se decidió analizar un conjunto de datos cuya información radica desde 1985 hasta 2015, es decir, se cuentan con datos oficiales de 30 años.

METODOLOGÍA

En primera instancia se recurrió a utilizar herramientas informáticas que pudiesen mostrar información puntual de los datos con los que se cuentan para hacer un análisis exploratorio de los mismos. Haciendo uso del lenguaje de programación R, junto con sus siguientes paquetes:

tidyverse, countrycode, rworldmap, ggplot2, gridExtra y broom.

Posteriormente se procedió a cargar el conjunto de datos a través de tidyverse, conociendo las variables con las que se cuenta, entre ellas: país, año, sexo, edad, número de suicidios, población, suicidios por cada 100 mil personas, país-año, IDH (Índice de Desarrollo Humano), PIB, PIB per cápita y generación, las cuales fueron analizadas con el objeto de erradicar inconsistencias. Se encontró que la columna IDH tiene dos tercios de valores nulos, por lo que se eliminó, así mismo se descartó la variable generación por tener incongruencias a través de los años, por último se dejó fuera a la columna país-año, al ser irrelevante en el análisis del conjunto de datos.

Por otra parte se localizó que 7 países cuentan con 3 años o menos de registros por lo que son eliminados dado que pueden ocasionar anomalías en el análisis, se eliminaron los datos de 2016 por un repertorio mínimo de datos, así mismo el continente se agregó al conjunto de datos utilizando el paquete countrycode y se descubrió que África tiene muy pocos países que proporcionan datos de suicidio.

Seguido de la limpieza de datos, se catalogaron en valores categóricos y ordinales y se dio

una visualización final del conjunto, procediendo a realizar un análisis global y particular de los datos, en donde se construyen escenarios para validar y desestimar hipótesis, con el objeto de idear marcos de referencia en la prevención de suicidios; en este punto paquetes como rworldmap, ggplot2, gridExtra y broom son los encargados del manejo de datos y construcción de recursos gráficos.

Uteriormente se crean modelos informáticos a través del lenguaje de programación python, con herramientas como son: pandas, numpy y sklearn quienes proporcionan el buen manejo de datos, funciones matemáticas (principalmente de álgebra lineal) y algoritmos de aprendizaje automático de manera correspondiente. Posteriormente se construye un modelo de predicción de suicidios en un contexto global el cual es entrenado con el presente conjunto de datos. Para finalizar se crea un modelo preliminar de perfiles suicidas a través de algoritmos de clasificación no supervisada que agrupa datos en grupos, basándose en sus características y teniendo como objetivo comprender el comportamiento del fenómeno dentro del contexto social.

CONCLUSIONES

El análisis de los datos aunado al uso de herramientas de ciencia de datos nos faculta a encontrar una estimación en los índices de suicidio por cada 100 mil personas en un contexto global.

El modelado de los datos nos permite tener una clasificación preliminar de los perfiles suicidas según indicadores de: edad, sexo, año, país y PIB per cápita.

Se cuenta con un análisis de tendencias globales, comportamiento del fenómeno por continente, país, sexo, edad, así como cifras mundiales distribuidas en mapas geográficos, cambio anual en las tasas de suicidio, las tendencias más pronunciadas y con mayor disminución, disparidad de género por continente y país, de la misma forma la disparidad de edad por continente, proporciones de suicidio masculino y femenino, por otra parte se cuenta con la correlación entre el PIB per cápita y el índice de suicidio, finalizando con las instancias de mayor riesgo en la historia y una comparativa de los índices de suicidio en México con países cultural, legal y económicamente similares.