



Análisis de la tasa de suicidios a nivel mundial

JORGE ANDRES MOGOTOCORO FAJARDO
JHEYSON ARLEY JAIMES ESTEBAN
ANDRES RICARDO HERNANDEZ TORRES



Motivación

¿Es importante tratar el tema del suicidio?



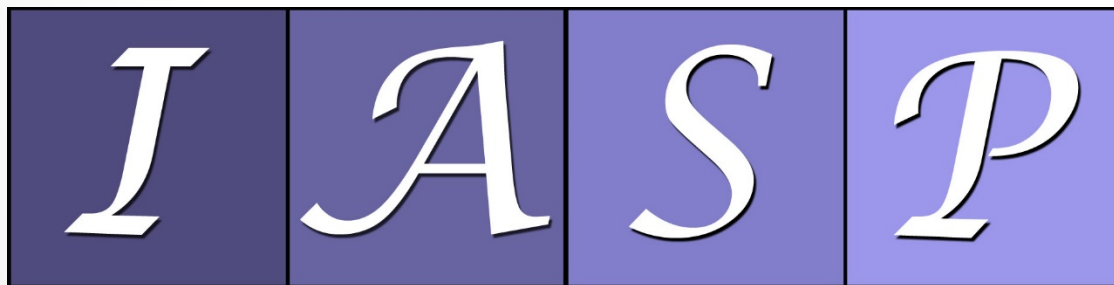
**Organización
Mundial de la Salud**

¿Según la OMS cerca de 3000 personas se suicidan cada día en el mundo y otras 60000 lo intentan, esto es equivale a un suicidio cada 30 segundos.

¿A quien le interesa?

IASP

(International Association for Suicide Prevention) Asociación internacional dedicada a la prevención del suicidio haciendo acompañamiento a las personas que lo deseen.



Línea 106

Línea habilitada en bogotá, totalmente gratis y 24/7, la cual escucha cualquier persona con problemas que piense en el suicidio como solución.

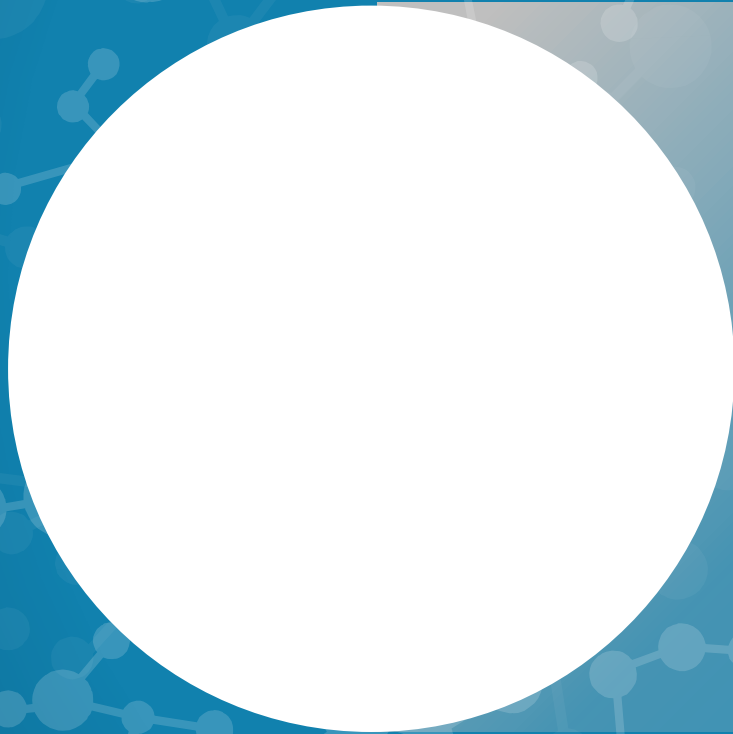




Objetivos

Objetivos

- Entrenar modelos de predicción que nos permitan visualizar un estimado de suicidios según características del entorno de un país.
- Identificar cuales son los factores más influyentes en la tasa de suicidios.
- Encontrar las mejores condiciones de un país para minimizar la tasa de suicidios.



Dataset

Dataset

El dataset describe el número de suicidios en 101 países desde 1985 hasta 2016 en grupos según el sexo y rango de edad.

	country	year	sex	age	# suicide...	# populat...	# suicide...	country...	# HDI for ...	# gdp_fo...	# gdp_pe...	generat...
1	Albania	1987	male	15-24 years	21	312900	6.71	Albania1987		2,156,624,900	796	Generation X
2	Albania	1987	male	35-54 years	16	308000	5.19	Albania1987		2,156,624,900	796	Silent
3	Albania	1987	female	15-24 years	14	289700	4.83	Albania1987		2,156,624,900	796	Generation X
4	Albania	1987	male	75+ years	1	21800	4.59	Albania1987		2,156,624,900	796	G.I. Generation
5	Albania	1987	male	25-34 years	9	274300	3.28	Albania1987		2,156,624,900	796	Boomers

Fuente: Kaggle

Tamaño Original: **27820**

Campos: **12**



Tratamiento de datos

Se realizó una limpieza del dataset cambiando todos los valores por números

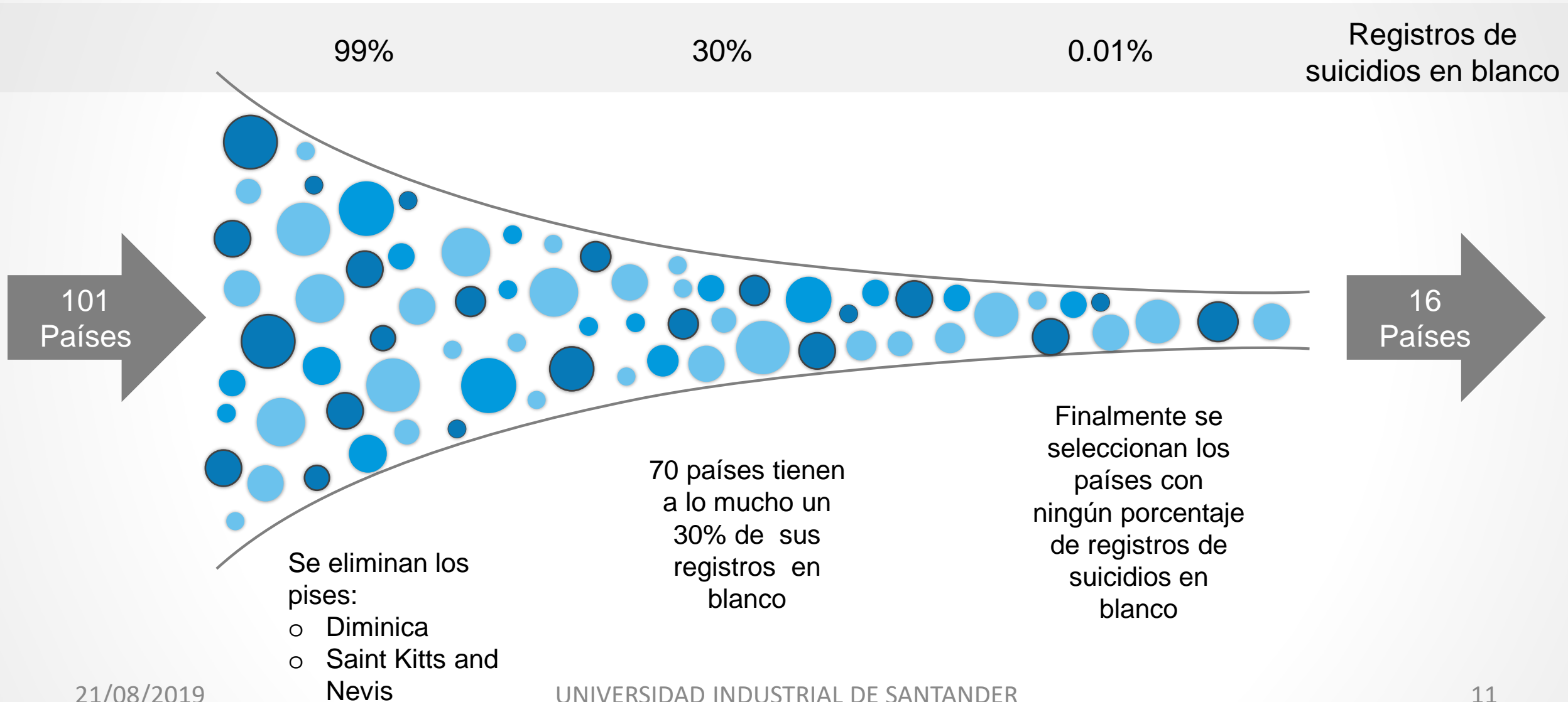
A = 1

Se llenaron los todos NaN de la columna HDI con el promedio de cada país.



Se elimino una columna con información redundante.

Análisis de valores de suicidios no registrados



Dataset luego del tratamiento

	country	year	sex	age	suicides_no	population	suicides/100k_pop	country-year	HDI_for_year	gdp_for_year	gdp_per_capita	generation
0	32	1985	0	5	202	363000	55.65	Argentina1985	0.694	88.417	3264	0
1	32	1985	0	4	485	1997000	24.29	Argentina1985	0.694	88.417	3264	0
2	32	1985	0	3	414	3346300	12.37	Argentina1985	0.694	88.417	3264	1
3	32	1985	1	4	210	2304000	9.11	Argentina1985	0.694	88.417	3264	0
4	32	1985	0	2	177	2234200	7.92	Argentina1985	0.694	88.417	3264	2

	country	year	sex	age	Suicides no	populati on	suicides/ 100k_po p	HDI_for_ year	gdp_for_ year	gdp_per_ capita	generatio n
count	<u>4920</u>	<u>4920</u>	<u>4920</u>	<u>4920</u>	<u>4920</u>	<u>4920</u>	<u>4920</u>	<u>4920</u>	<u>4920</u>	<u>4920</u>	<u>4920</u>
min	32	1985	0	0	<u>1</u>	4.0e+04	0.06	0.576	5.978	428	0
max	860	2015	1	5	11767	4.38052 1e+07	125.22	0.916	18120.7 14	60387	5



Selección de características

Selección de características



X

- Country
- Year
- Sex
- Age
- Population
- HDI for year
- GDP for year
- GDP per capita
- Generation

y

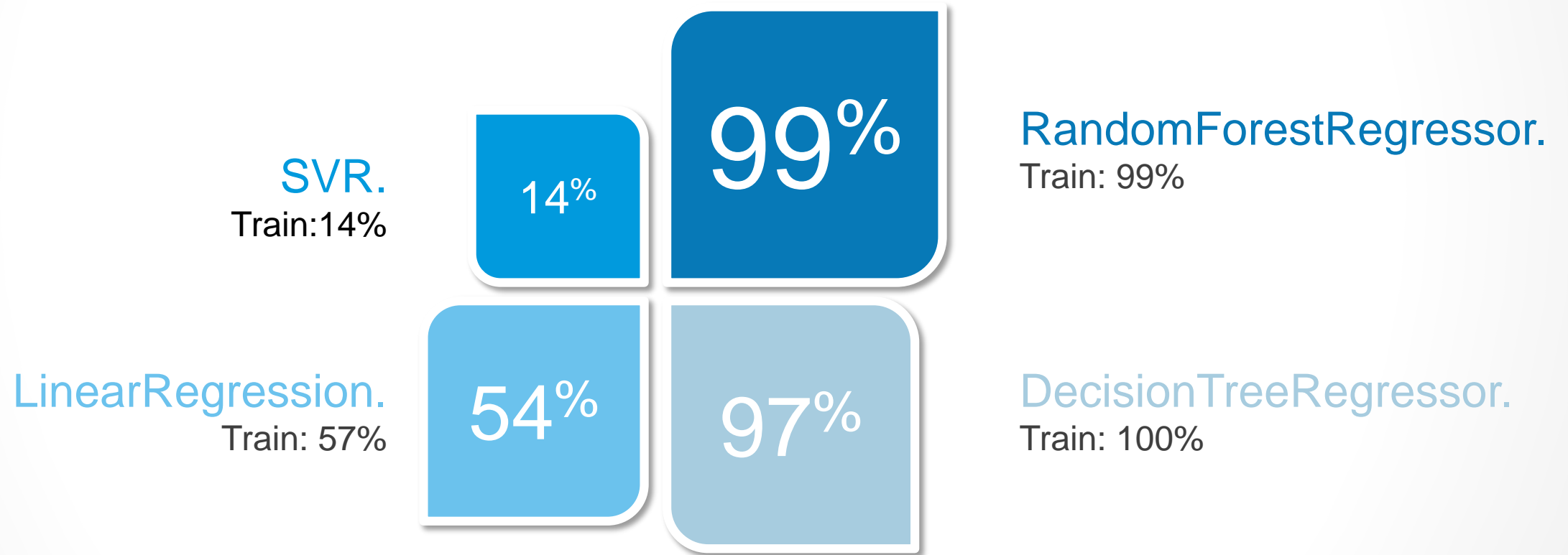
- No suicides



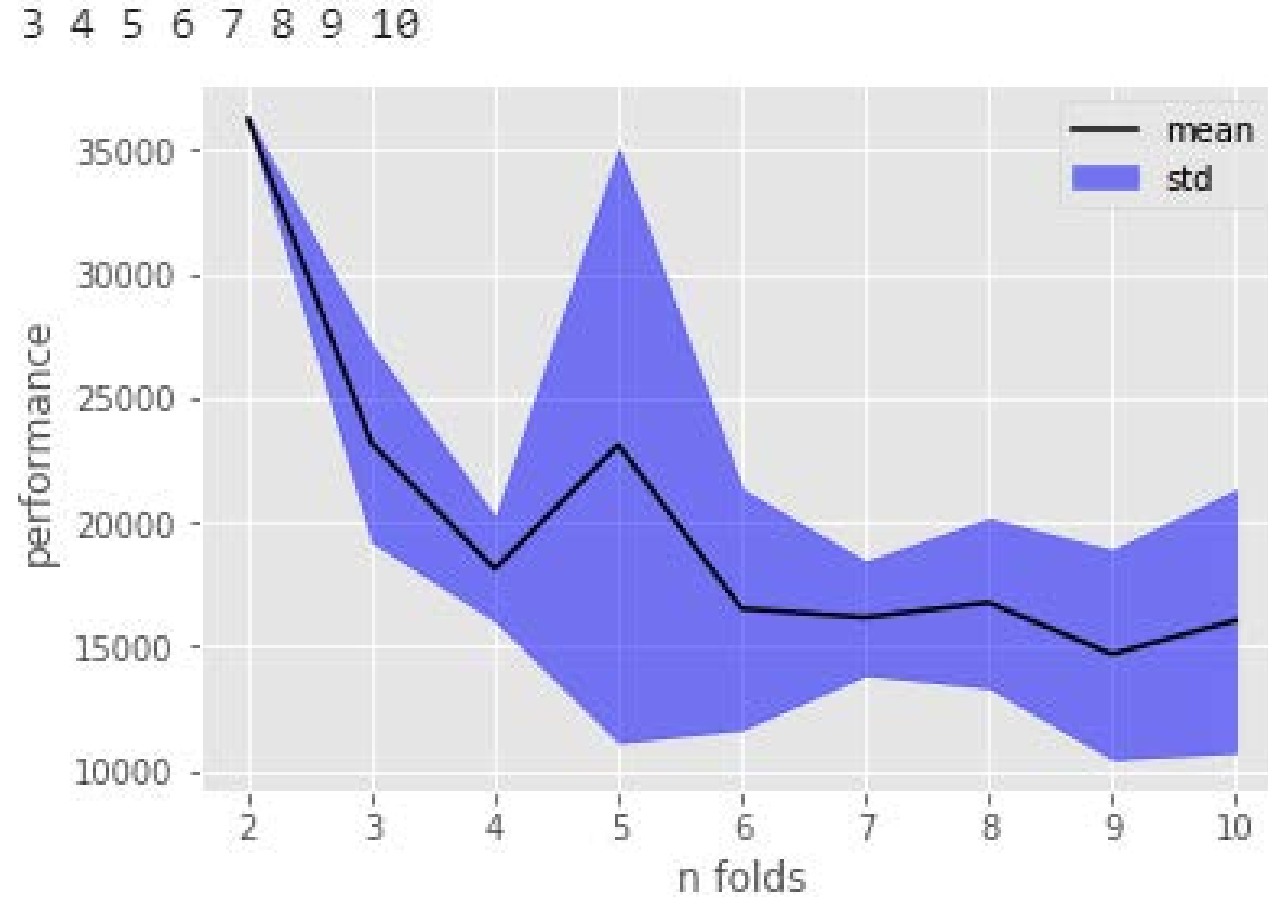
Modelos de regresión

Modelos de regresión

Modelo seleccionado: RandomForestRegressor

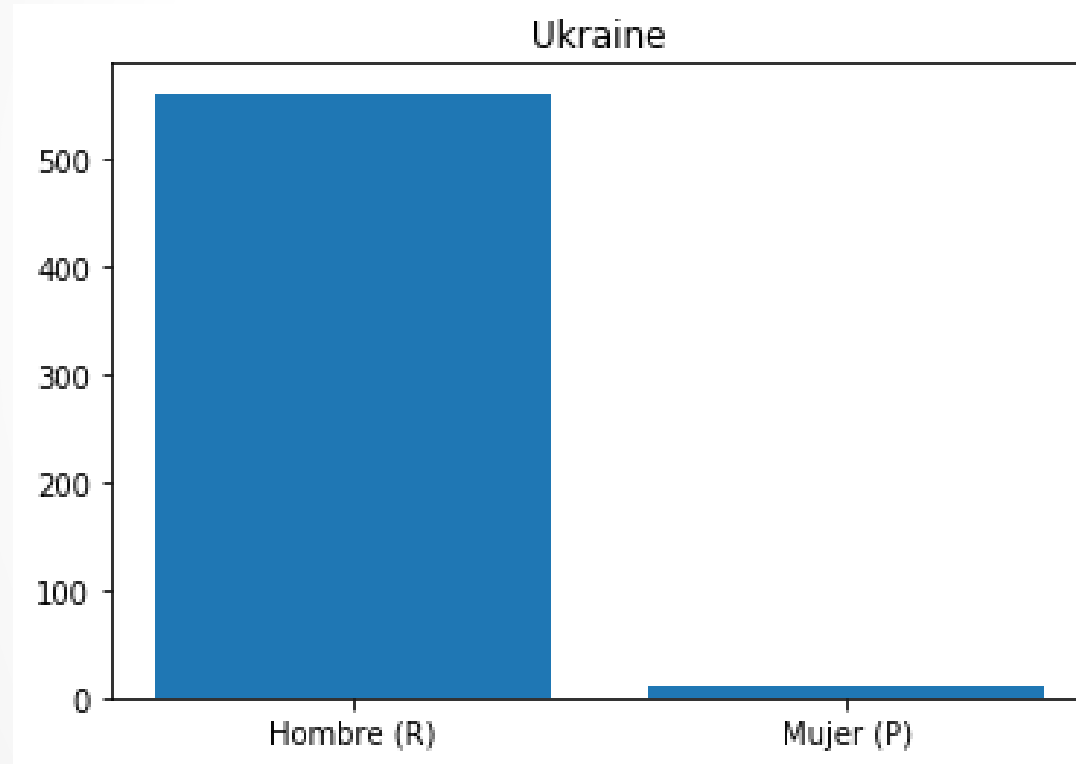


Evaluación del algoritmo usando una curva de aprendizaje

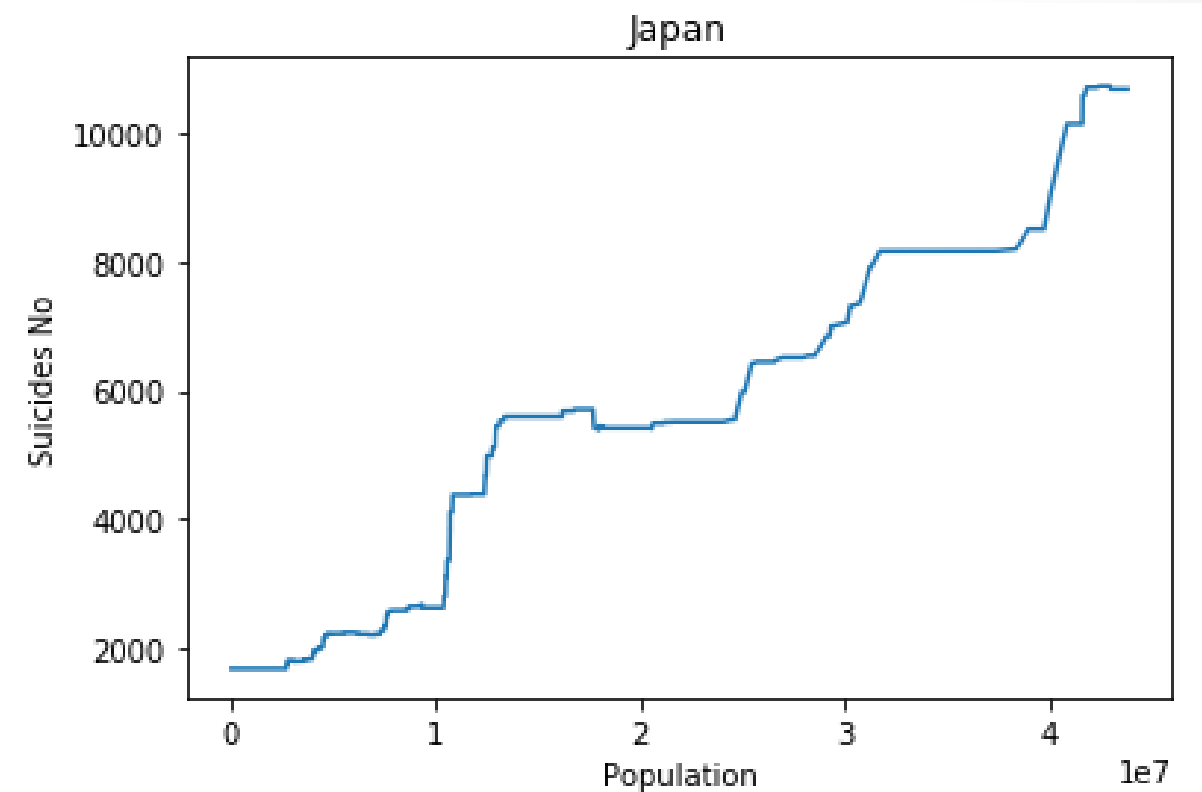


Análisis del modelo

Variación del sexo

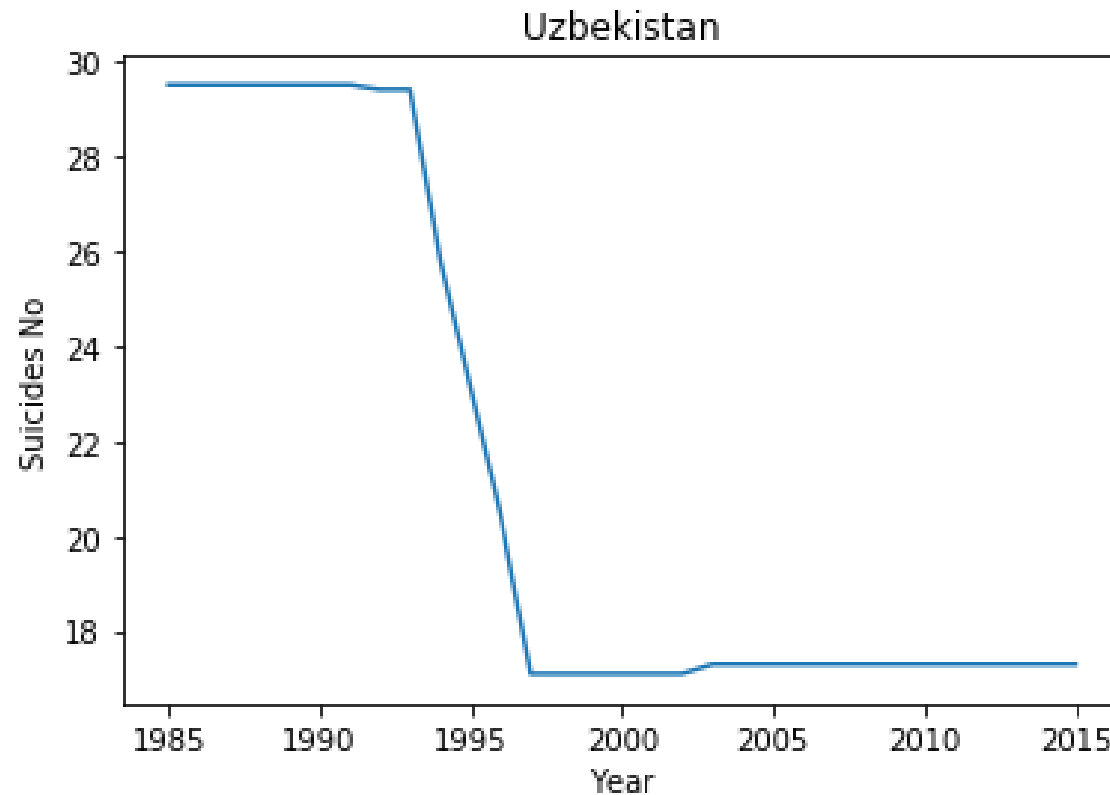


Variación de la población

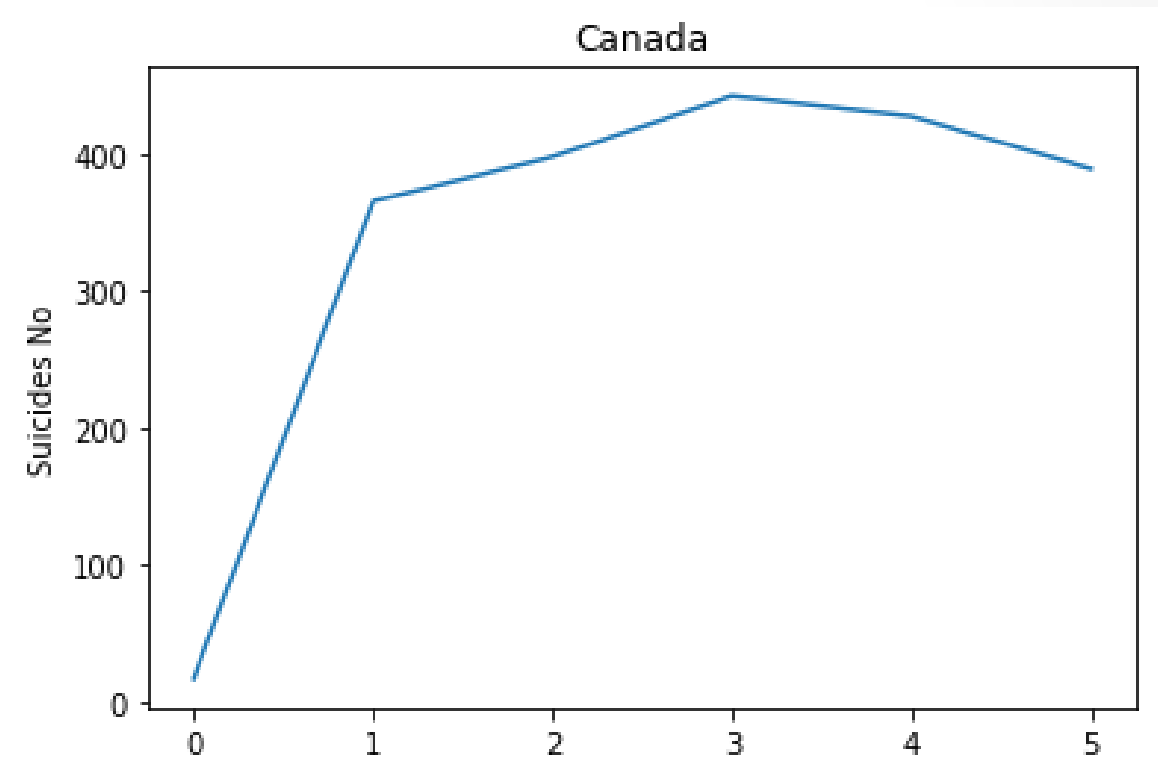


Análisis del modelo

Variación del año



Variación de la edad



Conclusión del análisis del modelo

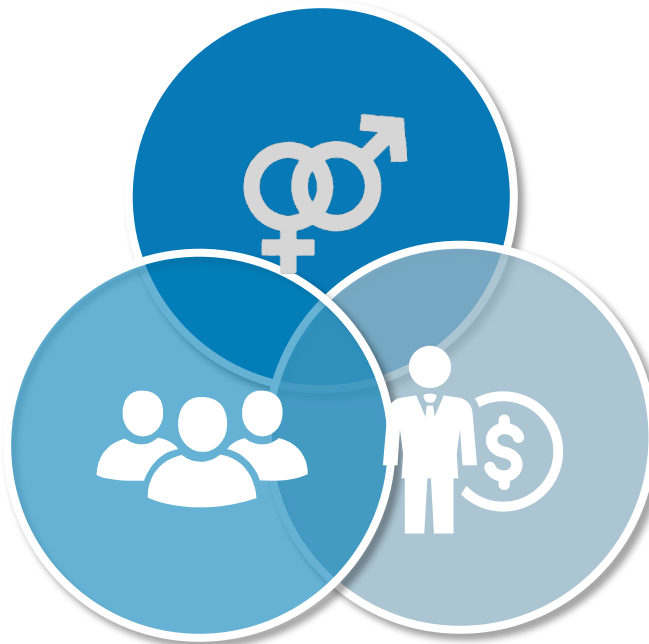
Sexo

Los **hombres** poseen una mayor tendencia al suicidio.

Al permutar el **país** en el modelo se encontró que tampoco era una característica relevante

Población.

A medida que el país tenga mas habitantes también lo es la cantidad de suicidios que pueden ocurrir.



Al permutar la columna del **año** no afecta el numero de suicidios lo cual impide usar el modelo para predecir tasas de suicidios en un futuro

Condiciones económicas y de calidad de vida.

Un país con mayor esperanza de vida, mas educado y con personas mas productivas no piensa en suicidarse



¿Cuál será la mejor configuración de características para disminuir la tasa de suicidios?

Algoritmo genético

Se realiza el proceso iterativamente hasta converger a una buena respuesta.

Función de inicialización
de la población

1

Parte 1

Le 14

Le 13

Parte 1

Le 14

Le 13

2

Función de reproducción

Función de mutación

3

Restricciones de la población



$$HDI = 0.134 * \log(GDP) - 0.55.$$



El **GDP for year** tiene una relación directa con **GDP per capita**

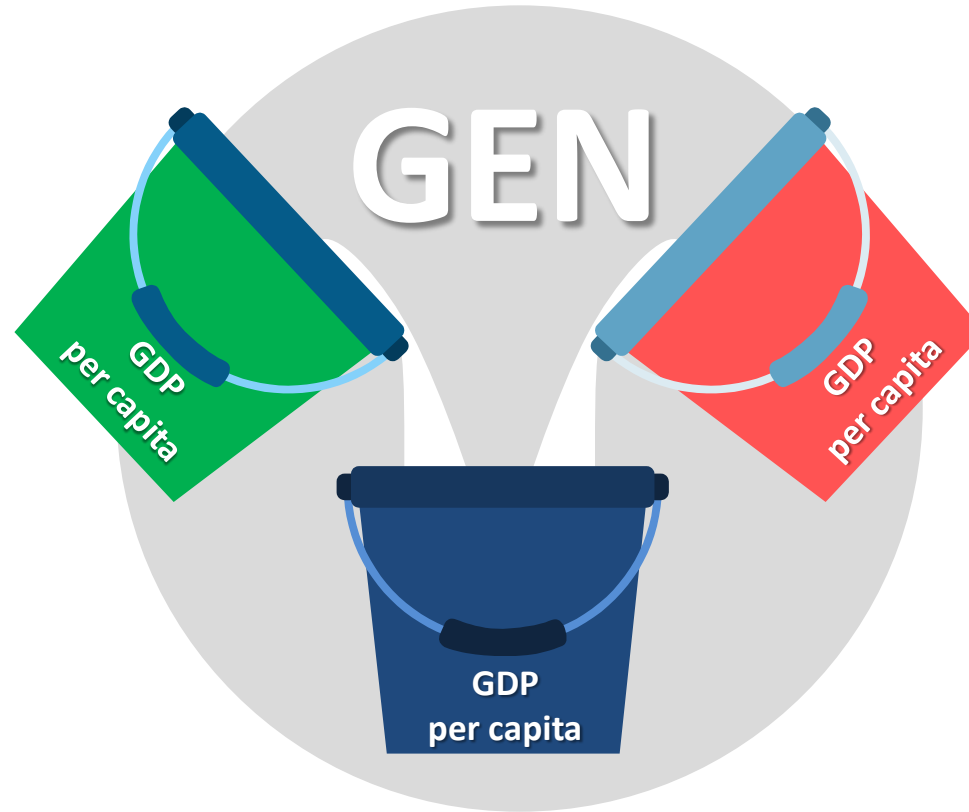


El valor de generación debe Concordar con alguna de las Edades propuestas en la solución.



Mutación

Aumento
En base a un porcentaje
de variación se **augmenta**



Disminución
En base a un porcentaje
de variación se **disminuye**

Individuo mutado

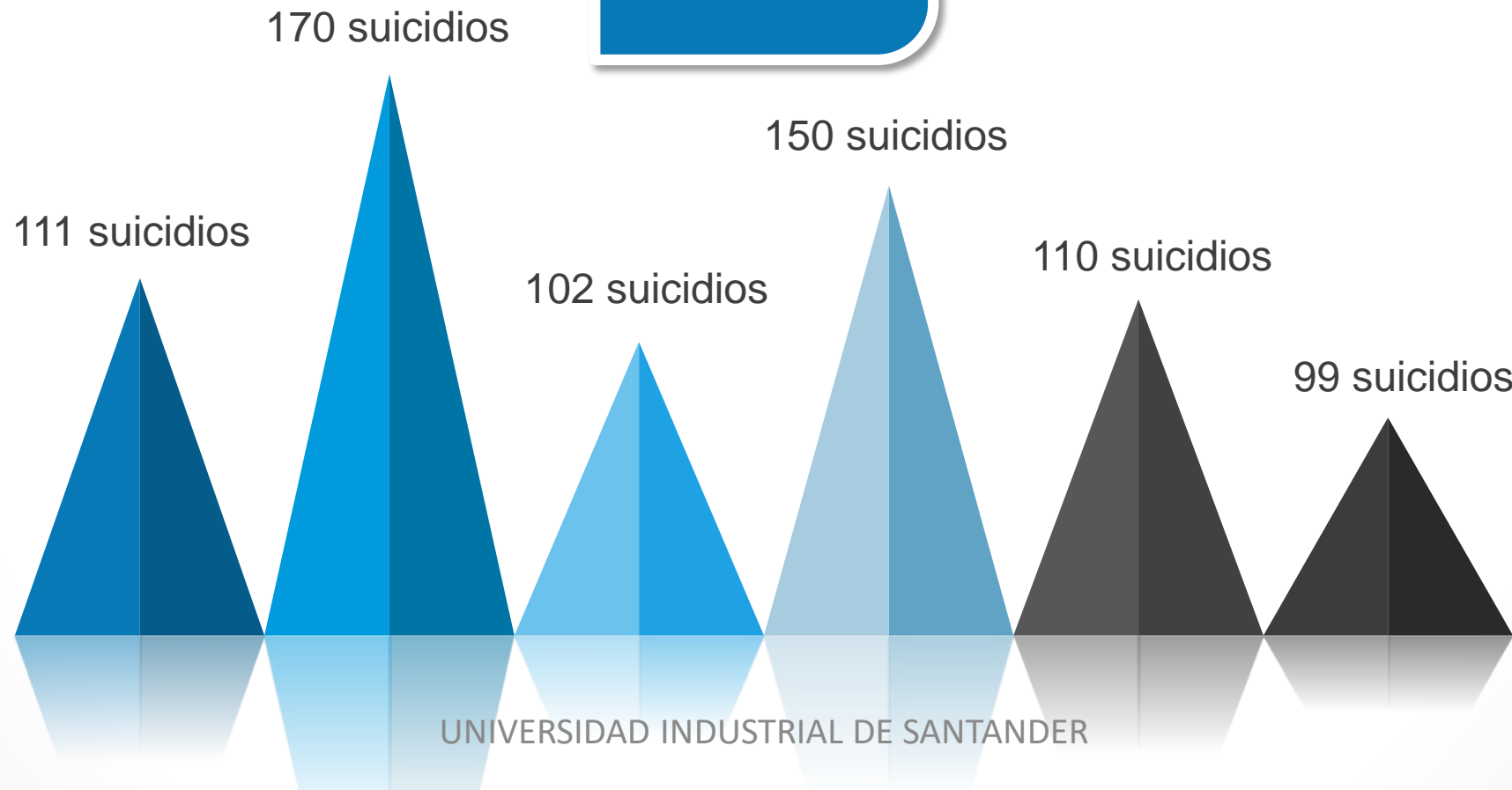
Gracias a las restricciones definidas anteriormente
no solo se esta variando el **GDP per capita**, el **HDI for year** y el **GDP for year** también
varian dado que son dependientes

Función de costo

99%

RandomForestRegressor.

Train: 99%



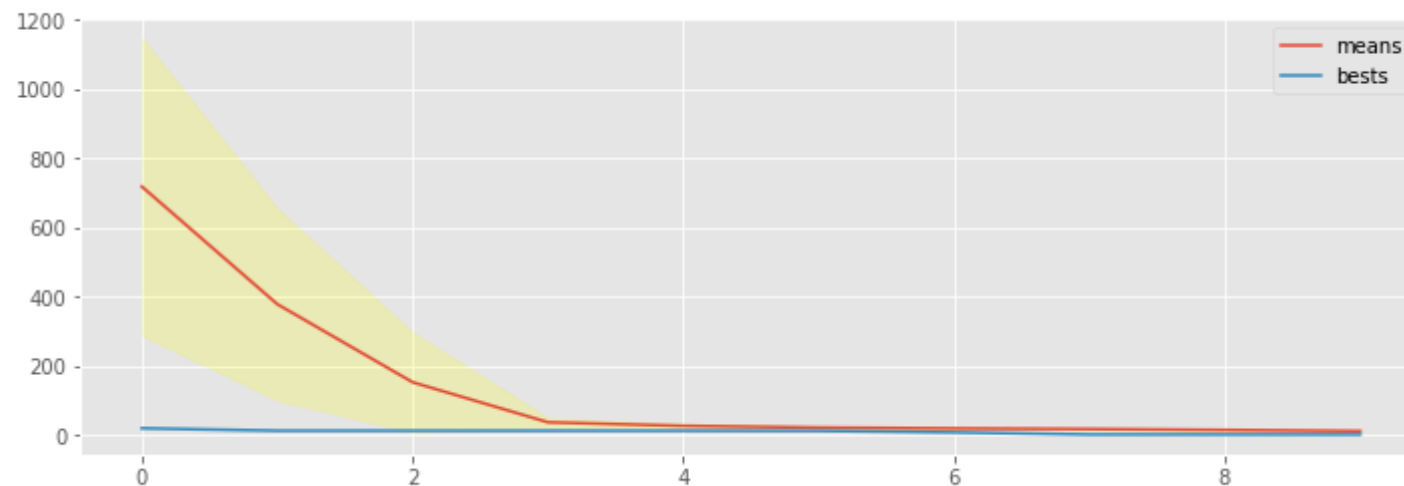


Resultados

country	Ecuador
year	1995
sex	Femenino
age	5-14 años
population	93657
HDI_for_year	0.80773819
gdp_for_year	736.38
gdp_per_capita	25144
generation	Millenials
no suicidios	2.4

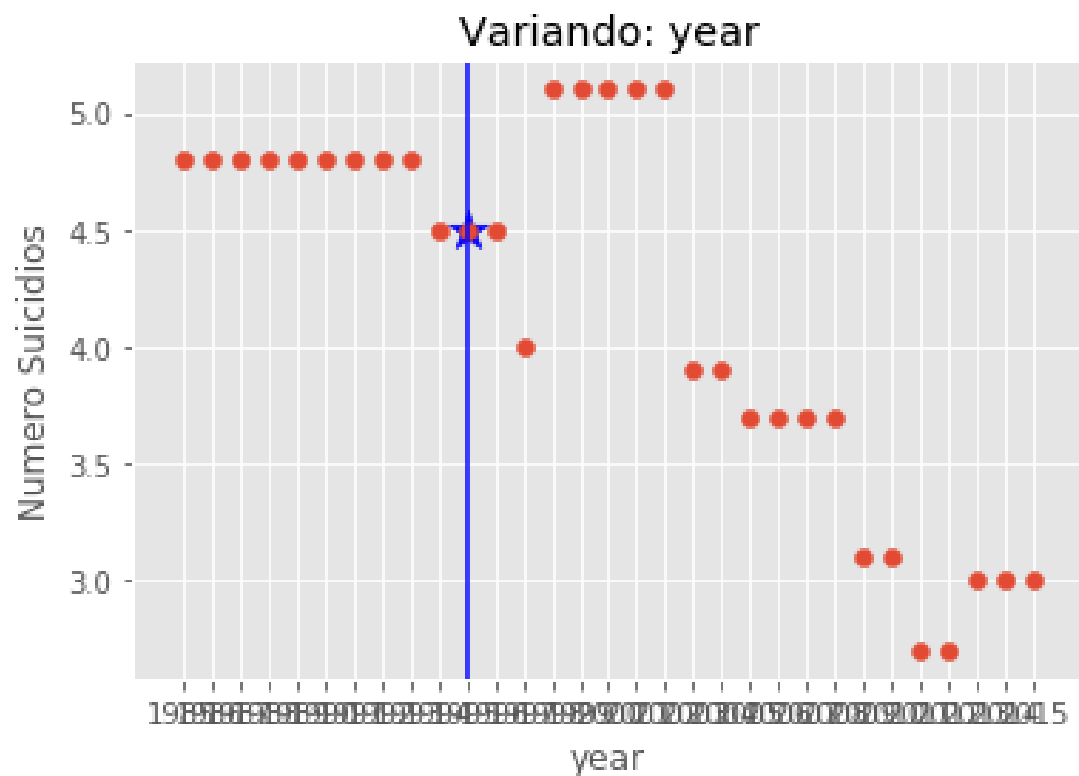
Mejor solución

HBox(children=(IntProgress(value=0, max=10), HTML(value=u'')))

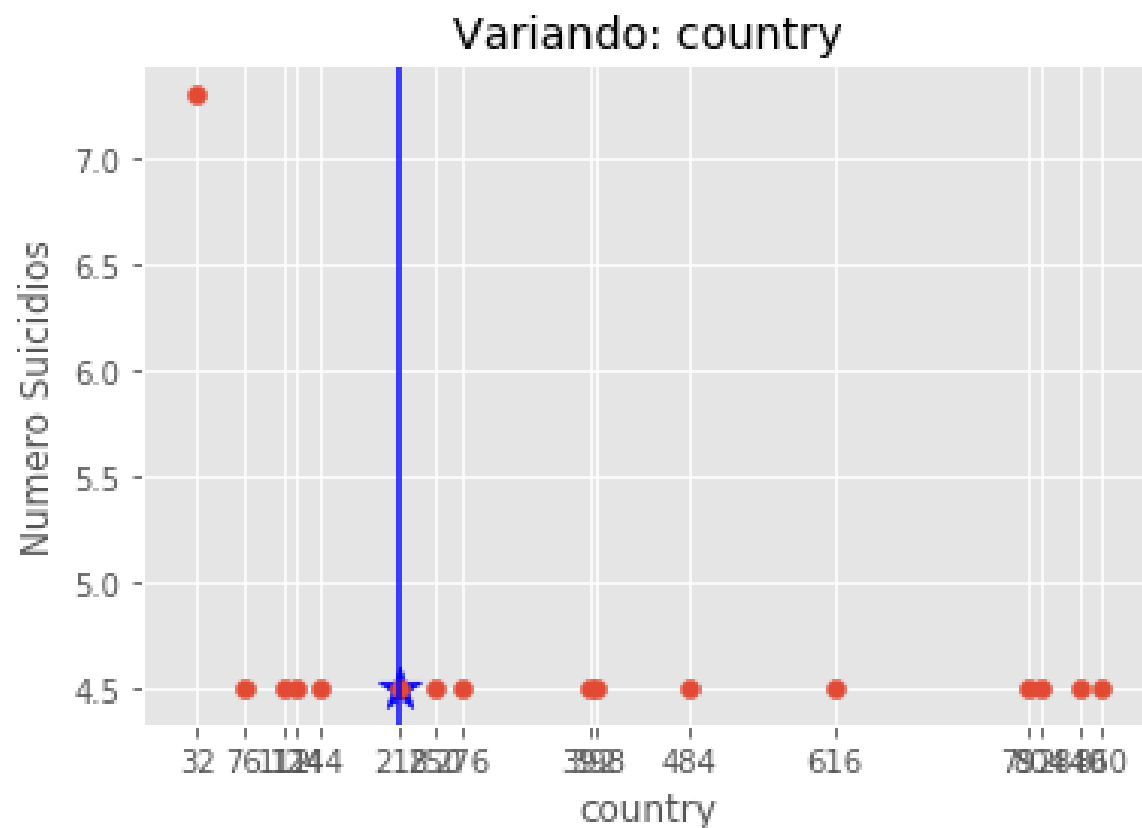


Análisis de la mejor solución

Año mínimo: 1985
Año máximo: 2015
Año individuo: 1995



32: Argentina
218: Ecuador

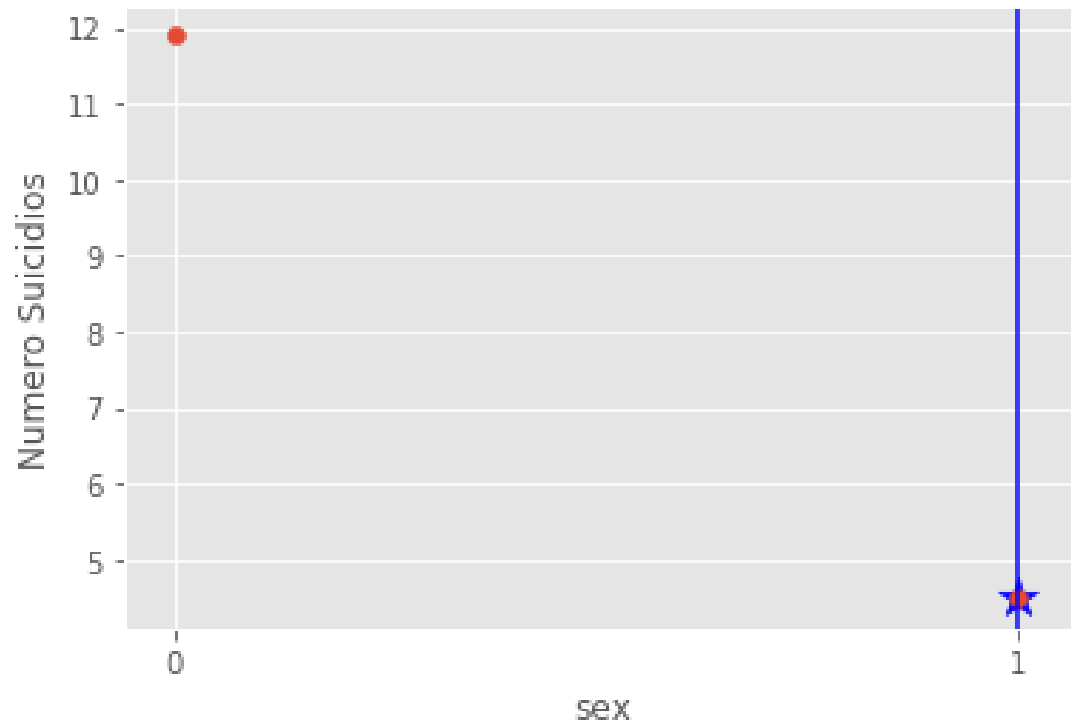


Análisis de la mejor solución

0: Hombre

1: Mujer

Variando: sex

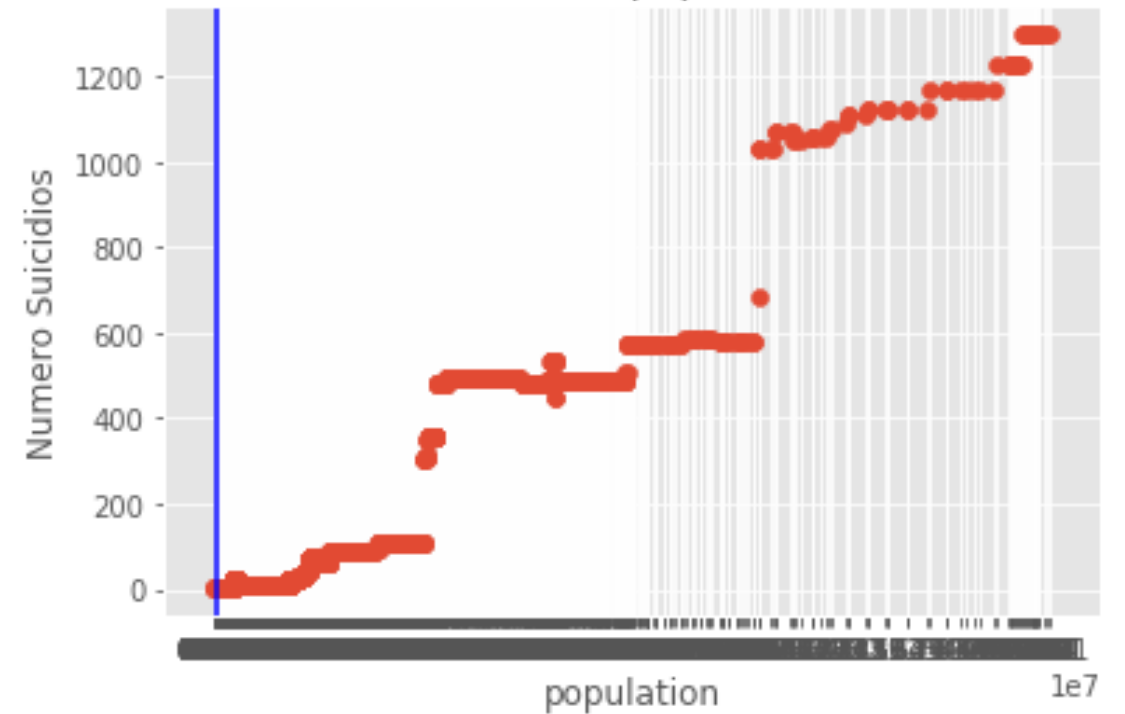


Pop minima: 40000

Pop maxima: 43805214

Pop individuo: 93657

Variando: population





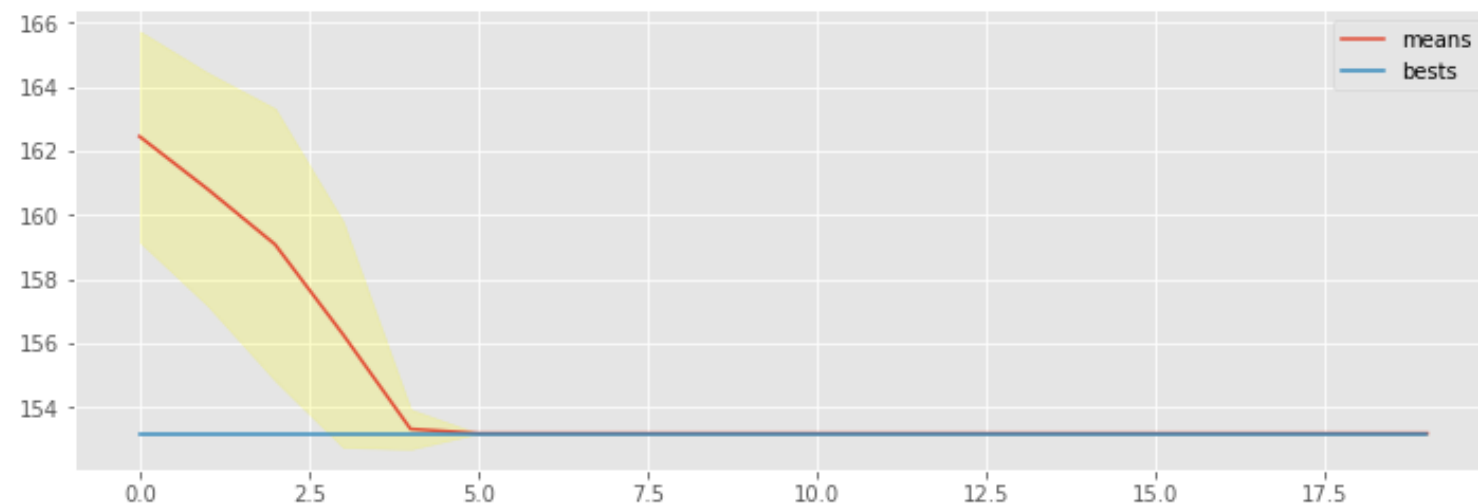
Mejores condiciones de vida para Colombia

country	Colombia
year	<u>2015</u>
sex	<u>Femenino</u>
age	<u>15 a 24 años</u>
population	<u>4070390</u>
HDI_for_year	0.630020
gdp_for_year	527.813
gdp_per_capit a	6675.0
generation	<u>Millenials</u>
no suicidios	153.2

Mejor solución

```
HBox(children=(IntProgress(value=0, max=20), HTML(value=u'')))
```

<Figure size 432x288 with 0 Axes>

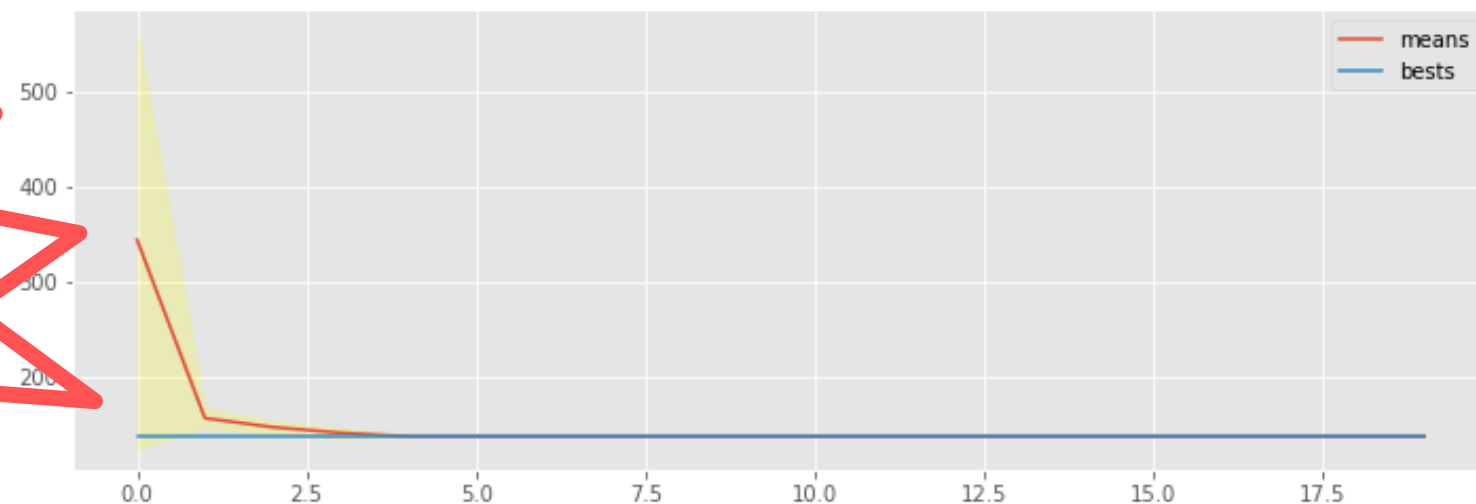


country	Colombia
year	<u>2015</u>
sex	<u>Femenino</u>
age	<u>15 a 24 años</u>
population	<u>4070390</u>
HDI_for_year	0.896866
gdp_for_year	1549.131
gdp_per_capit a	48809.0
generation	<u>Millenials</u>
no suicidios	137.5

Mejor solución

HBox(children=(IntProgress(value=0, max=20), HTML(value=u'')))

<Figure size 432x288 with 0 Axes>



153.2



Conclusiones

Conclusiones



Las características que mas repercuten en el numero de suicidios fueron: **Sexo, población, HDI for year, GDP per capita y GDP for year.**



Si generamos un aumento respecto a características como las referentes a la condición de vida y economía del país, la tasa de suicidios efectivamente se reduce.



Fue eficiente proponer un modelo de regresión para poder calcular el numero de suicidios ya que obtuvimos una alta precisión además también cumplió la tarea de función de costo en el algoritmo genético.



Generar características de condición humana y económica es una tarea compleja que en nuestra implementación presenta limites ya que las condiciones ideales para minimizar una tasa de suicidios en un país en especifico en muchos casos son prácticamente imposibles de ejecutar para una nación.

GRACIAS POR LA
ATENCIÓN PRESTADA