

Data Science for Business Good: Stopping suicides

Haerim Yun

Azamat Zhangeldin

Emili Vilanova

Louis David

Fabio Fruchtengarten

Yaser Farook

Hubert Burakiewicz



Suicide is a very serious problem being on the top of WHO agenda



approximately

800,000

people dying by

suicide

every year

one death every 40 seconds



Suicide is 100% preventable

23

people will commit **suicide**by the end of this presentation (**45%** of our class)

20



unsuccessful suicide attempts per 1 death by suicide

18

<u>.Ш.</u>

worldwide cause of **death**

More deaths
from suicide
than from
homicide and
war together





Our goal and objective is driven by WHO global agenda



Our Client



Dr Ren Minghui

Assistant Director-General for Communicable Diseases

- member of WHO global leadership
- overseas a complex portfolio of technical programs covering, inter alia, mental health and substance use
- Leading global actions and suicide prevention programs

Objectives and goals

WHO goal: reduce suicide rates by 10% (WHO Mental Action Plan 2013-2020)
UN goal: decrease suicide rates by 1/3 (UN Sustainable Development Goals 2030)

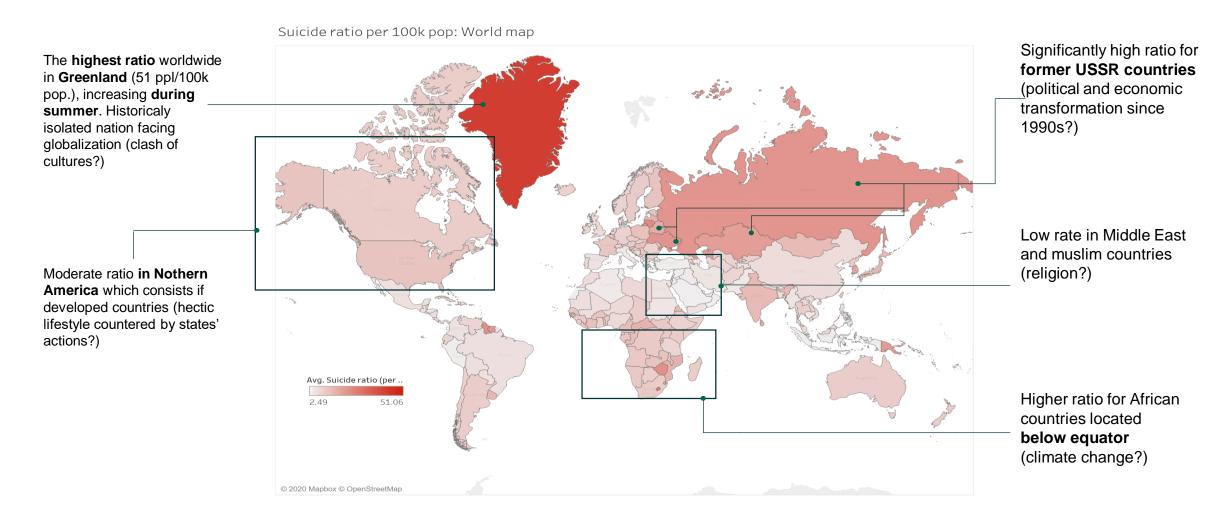
Our role and help:

- Assessing the goals of WHO and UN (feasibility analysis)
- Understanding the drivers of suicides
- Recommend actions/ initatives for WHO



Suicide ratio differs significantly among regions, with Asia reporting the highest rate



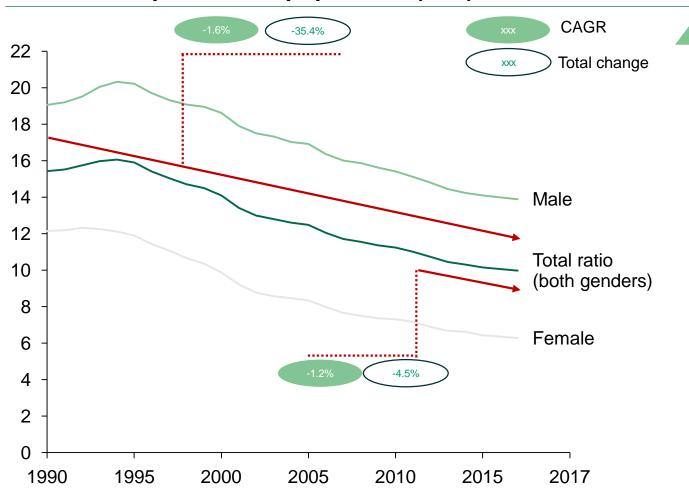




Over the last 28 years the suicide ratio has been steadily decreasing from 15.4 in 1990 to 10.0 in 2017



Suicide rate per 100k of population (unit)



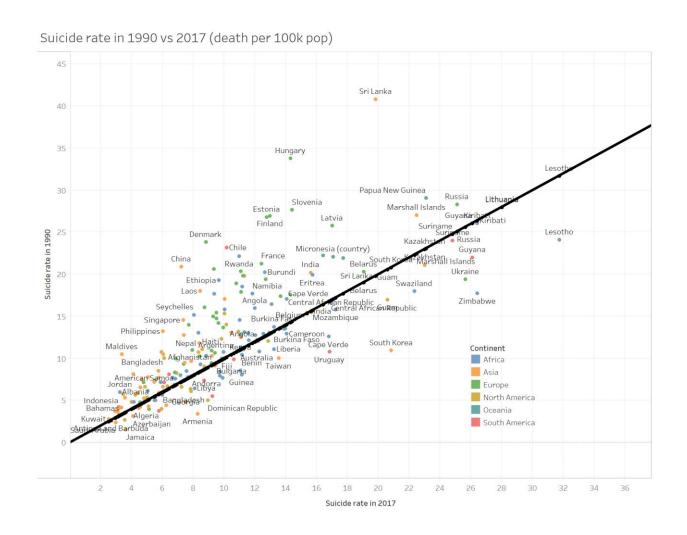
Take-aways

- Total suicide ratio is steadily declining from 15.4 in 1990 to 10.0 in 2017 (-1.6% CAGR)
- Dynamics of decline is going down which is a negative trend (could reverse)
- WHO target to decrease suicide ratio by 10% from 2013 to 2020 seems to be unrealistic (2017 vs. 2013 decline in ratio of 4.5%)
- Suicide rate is significantly higher for men than for women (two times difference)



Majority of countries report declining suicide rate, however, the trend varies among regions





Take-aways:

- Majority fo countries reporting decrese in suicide ratio in 2017 vs. 1990 (lying above the grey line):
 - every European country (except from Ukraine) with no change or decrease in suicide ratio
 - strong declining trend for Asian countries (esp. for Sri Lanka)
- Varying trend for Oceania and Africa
- Highest increases for African and South American countries



The leveraged process was complex and comprehensive



Process flow

Collecting and cleaning data

Time-series analysis

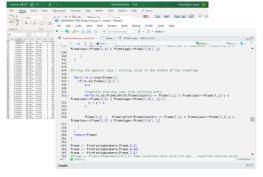
Regression models

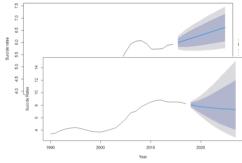
Drawing insights and developing recommendations

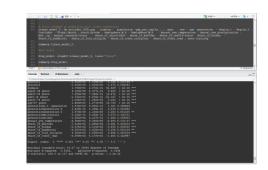
Description

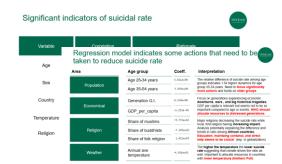
- Wide research and data gathering and inital analysis
- Cleaning (fixing missing values and • removing)
- Determining timeseries models
- Predicing suicide ratio in the future
 - Analyzing trends per regions
- Building regression models
- Data featuring
- Choosing the best model
- Data analysis
- Developing recommendations for WHO

Outcome











We complemented WHO's data to form a comprehensive dataset INSEAD

Area	Variable	Unit	Describtion	
Dependent variable	suicides.100k.pop	num.	Ratio of suicides per 100k of population	
Year	Year	num.	Year of observation	
	Country	unit	Name for UN recognized country	
	Region.1	unit	Tier-1 subregion as per UN classification	
	Region.2	unit	Tier-2 subregion as per UN classification	
Geographical	Continent	unit	Continent as per UN classification	
	Hemisphere_N_S	unit	Hemisphere where country lies (Northern or Southern)	
	Hemisphere_W_E	unit	Hemisphere where country lies (Western or Eastern)	
	sex	unit	Gender of suicide victim	
	Age	num.	Age group of suicides	
Population	suicides_no	num.	Total number of suicides for given observation	
	Population	num.	Total population of country	
	Generation	unit	Type of victim generation (e.g., generation X, generation Y)	
Economical	gdp_for_year	USD	GDP for given observation	
	gdp_per_capita	USD	GDP per capita for given observation	
	Global.NorthSouth.divide	N/A	Assignment of Global North-South division (ecopol. segmentation)	
	Share_of_christians	%	Share of Christians for given observation	
	Share_of_muslims	%	Share of Muslims for given observation	
	Share_of_unaffiliated	%	Share of unaffiliated for given observation	
Religious	Share_of_hindus	%	Share of Hindus for given observation	
rteligious	Share_of_Buddhists	%	Share of Buddhists for given observation	
	Share_of_folk_religion	%	Share of folk religious followers for given observation	
	Share_of_other_religious	%	Share of other religious followers for given observation	
	Share_of_other_jews	%	Share of Jews for given observation	
Weather	Annual_ave_temperature	Cel. degrees	Average yearly temp. In 1961-1990	
	Annual_ave_precipitation	mm	Ave. annual precipitation for country	
	Annual.sunshine.hours	hours	Ave. annual sunshine hour	
Others	AveIQ	unit	Ave. IQ for given observation World Health	
orce: Team Analysis Organization				

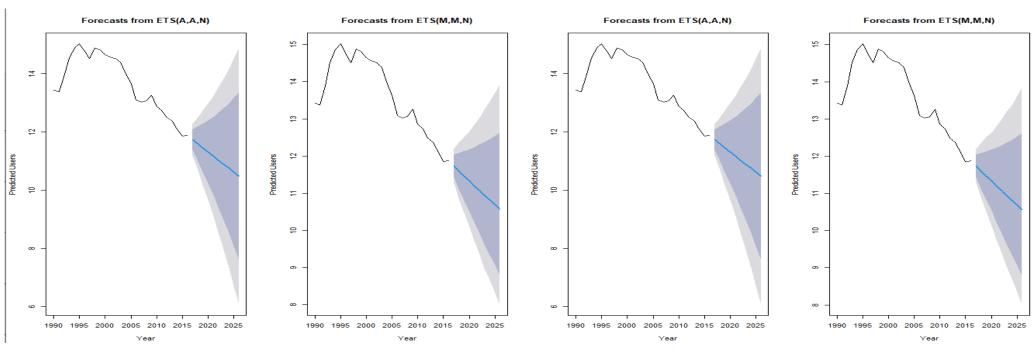
Source: Team Analysis

Time-series modelling for the world dataset



Insights

- Multiplicative smoothing ETS method (M,M,N) has the least error of the different TS models.
- Globally, the number of suicide per 100K population is forecasted to fall gradually.
- By 2030, we will have 22% less suicides than today for a given set of population. Risk of not achieving UN Sustainable Development Goals for 2030 with the current trend (Reduce the suicide rate by 33%).
- Each country has a unique trend in suicide rates and the global dataset averages out these individual trends hence providing little insight in our analysis.
- In the next few slides, we analysis granularly this data at continent and country level.



Source: WHO, Team Analysis

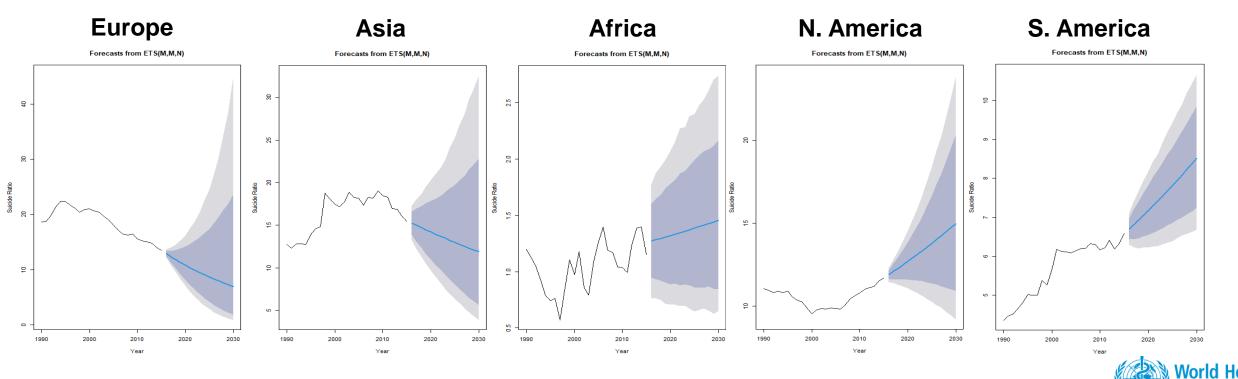


Forecasting using time-series model for the continents



Insights

- Time series analysis of the **5 populous continents** using the more accurate **(M,M,N) ETS model** gave better insights in the trends.
- We can see the suicide rates per 100k increasing in Africa, North America and South America, whereas the rates are
 decreasing in Europe and Asia. Possible factor could be increasing economic growth and prosperity in Asia and
 higher standards of living in Europe contributing to this trend. Further regression analysis could provide better
 understanding of the key factors.



Source: WHO, Team Analysis

Time-series: Country Level Predictions



Mexico	Armenia
Mexico	Armeni

Models Considered

AAN, MMN, AAA, MMM, TBATS

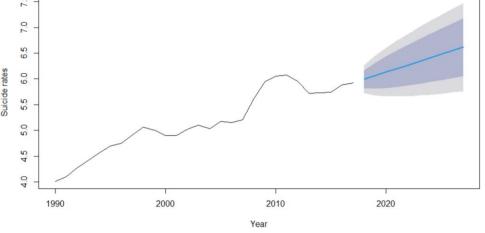
AAN, MMN, AAA, MMM, TBATS

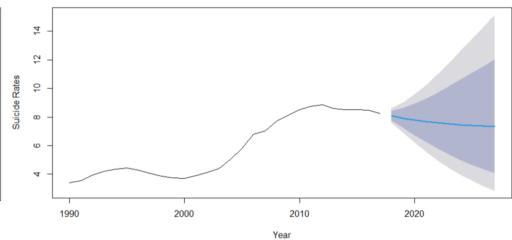
Model with least mean error

AAN Error of 0.918 TBATS Error of 1.274

Plots of selected Model

(10yr prediction after 2017; Confidence interval 0.8, 0.95)





interpretation

The plot shows that the suicide rates of Mexico would steadily increase over 10 years, amounting to 6.62 in 2017 from 5.92 in 2007. This indicates the main drivers of suicide should be contained while setting up a goal of keeping the rate in 2017 below 5.76(low 0.95 estimate)

The plot shows that the suicide rates of Armenia would slowly decrease over 10 years, amounting to 7.35 in 2017 from 8.26 in 2007. Although the rate is decreasing yet it is still higher than that of Mexico. This indicates that Armenia should set up an ambitious goal for 2017 below 4.08 (low 0.8 estimate)



Linear regression model



Selected Model

Independent variable: suicides/100k population

suicides.100k.pop ~ country + population + gdp_per_capita.... + year + sex + age +
generation + Annual_ave_temperature + Share_of_muslims + Share_of_hindus +
Share_of_buddhists + Share_of_folk_religion + Share_of_other_jews

R²=0.53, some heteroskedasticity

Data featuring

1) Removed variables with low significance, after 1st run of the model

population, suicide_no, HDI index, gdp_for_year...., Region.1, Region.2, Continent, Global.North...South.divide, Hemisphere_N_S, Hemisphere_W_E, Annual_ave_precipitation, Ave._IQ, Annual.sunshine.hours

2) Segment GDP per capita and add lagged variable for suicides ratio (year -1)

High impact on R², but loss of significance of independent variables



Significant indicators of suicidal rate



Variable	Correlation	Rationale		
Age Positive		Increasing age, increases the number of deaths. E.g. The coefficient for "+75 years" is 4.5x that of "25-34 years"		
Sex	Positive	Men are more likely to commit suicides than women. Men suicides are +17 every 100k people than women		
Country	Mixed	The country of origin is a strong indicator of suicidal rate. The coefficient for Canada negative, indicating people are less likely to commit suicide than Mexico, which has a positive coefficient		
Temperature	Negative	Average temperature is negatively correlated with suicide rates		
Religion Source: Team Analysis	Mixed	Hindus and Buddhists are less likely to commit suicide than other categories of population World Health Organization		

Regression model indicates some actions that need to be taken to reduce suicide rate

Area	Age group	Coeff.	Interpretation	
	Age 25-34 years	3.611e+00	The relative difference of suicide rate among age groups indicates 1.5x higher dynamics for age group 35-54 years. Need to focus significantly more actions and funds on older groups	
Population	Age 35-54 years	5.970e+00		
Economical	Generation G.I.	6.258e+00	Focus on generations experiencing economic downturns, wars, and big historical tragedies.	
LCOHOIIICai	GDP_per_capita	-1.221e-04	GDP per capita is relevant but seems not to be so important compared to age or events. WHO should allocate resources to distressed generations	
	Share of muslims	-9.755e+02	Major religions decreasing the suicide rate while local, fold religion having increasing impact. Analysis potentially explaining the difference and trends in ratio among African countries. Education, maintaing contanct, and direct help seems to be crucial (esp. in globalization)	
Religion	Share of buddhists	-7.495e+02		
	Share of folk religion	1.012e+03		
Weather Annual ave temperature		-4.561e+01	The higher the temperature the lower suicide rate suggesting that climate drives the ratio as well. Important to allocate resources in countries with lower temperature (Nothern Poll)	

Source: Team Analysis

Take aways, recommendations at global level and next steps



Key take aways

- Declining trend of suicide rate on global level (incl. predictions) that varies among the continents. UN and WHO targets seem not to be realistic to meet (not achieving goals as per time series model)
 - strong declining trend for Europe and Asia
 - Significant increasing trend for Africa
- The main drivers of suicide ratio:
 - ∘ Age **↑**
 - Sex ★
 - Country
 - Temperature
 - Religion

Selected recommendations

- Increase funds/ efficiency of spending money on fighting against suicide ratio. Current means not sufficient to achieve goals
- Increase activities esp. in Africa (least developed countries)
- Allocate resources:
 - for older people, e.g., psychological support in nursing home care, financial support
 - towards males, e.g., combine substance abuse therapies with suicide prevention programs
 - country specific, e.g.
 educate and offer direct help

Next steps

- Evaluate efficiency of WHO efforts and programs
- Run regression model to forecast suicide ratio for the next 2-3 decades to validate time-series model
- Gather detailed data on country level esp. in Africa to understand deeper drivers of suicide rate
- [TBD] Segment countries based on their characteristic to tailor programs and efforts



