VISUAL ANALYTICS

Project: "Tobacco: a silent killer"

Authors:

Emir Nurmatbekov

Manh Hung Nguyen

Julio Candela

Supervisor:

Petra Isenberg

January 2021

Table of contents

1. Team members and contributions

Member	Contributions
Manh Hung Nguyen	- Smoking deaths tab - Navigation panel - Deploy first version to Heroku - Report
Emir Nurmatbekov	- Tobacco consumption tab - Report
Julio Candela	- Home Tab - Tobacco Control Tab - Report

2. Introduction

The Tobacco epidemic is considered one of the biggest public health threats by the World Health Organization (WHO). WHO estimates millions of deaths per year caused by the frequent use of Tobacco.

However, WHO also states that Tobacco addiction is preventable and many measures are being applied to reduce its consumption as a world priority.

Key facts

- 20% of the world's population smoke tobacco everyday. WHO consider Tobacco as an epidemic
- 7 million people a year died as a result of Tobacco Smoking (more than Covid-19)
- Only 40% of the countries efficiently monitor this issue.

Purposes and Audiences

The visualisations provided in the project address the issue of smoking not as an individual problem but as a global issue. The tool was made for authorities or anyone who has access to authorities and wants to show the seriousness of tobacco consumption to them who in their place, can take actions to change the situation. The structure of the page is built in a consecutive manner, where the user sees first the scale of the problem, second, the consequences, and third, what measures were/are being taken. Such a narrative allows the user to build a convincing presentation/story to influence people in power.

Motivation

As stated by the World Health Organization, Smoking Tobacco is considered a harmful epidemic killing approximately 7 million people each year. While most of the deaths are the result of direct smoking, there is still 15% of deaths which result from passive smoking. Therefore, smoking is a problem that involves everyone since it affects the public health. With this purpose, the main motivation of our analysis is to expose the alarming scale and risks of Tobacco smoking. Likewise, we were encouraged to discover which countries were taking this issue seriously and force them to implement measures to prevent Tobacco consumption.

Tools

The charts shown in our website are built using mainly Altair and Streamlit in Python. The first version of the website is deployed to Heroku server. However, it did not run as smoothly as we wanted, which does not provide a good experience to users. In the final version, we switched to using Streamlit sharing feature which runs directly on top of our repository. The changes are updated automatically after every commit. There are some additional libraries that we used, such as, pandas, pyplot, numpy, json, etc.,. The source code and the data are stored in our github repository.

3. Feature Descriptions

Tobacco Consumption

The first feature which the user sees on the Tobacco Consumption page is country selection. It is a very intuitive select box where one can select a single or multiple countries of interest. Focusing on specific countries is the main element of an analysis of a global issue such as smoking. The box allows the user to select "World" too; in that case, one can compare a country's tobacco consumption trend to the rest of the world.
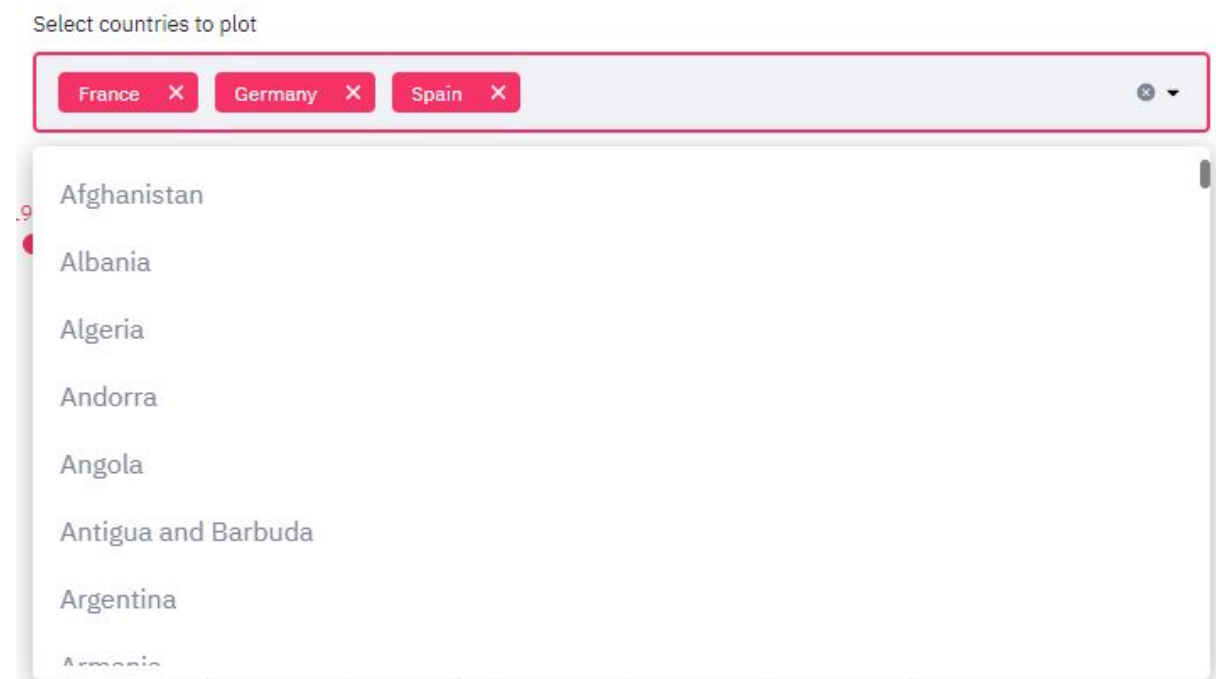
Year slider is the second main tool that can be used to zoom into a specific time period and carefully analyse and compare values for different countries.

After the countries and the time period have been selected, the tobacco consumption tab presents two charts: "Share of adults who smoke every day" and "Consumption per smoker per day". Both of the charts show the data for the chosen countries and the time period. Countries are encoded with color and the legends are located above the charts. There are also two synchronised rulers which help the user to see the exact values of the y-axis in both charts. The rulers are very helpful when drawing connections between the % of every-day-smoking adults and the number of cigarettes consumed per day per smoker.

Share of adults who smoke every day(%)

Consumption per smoker per day
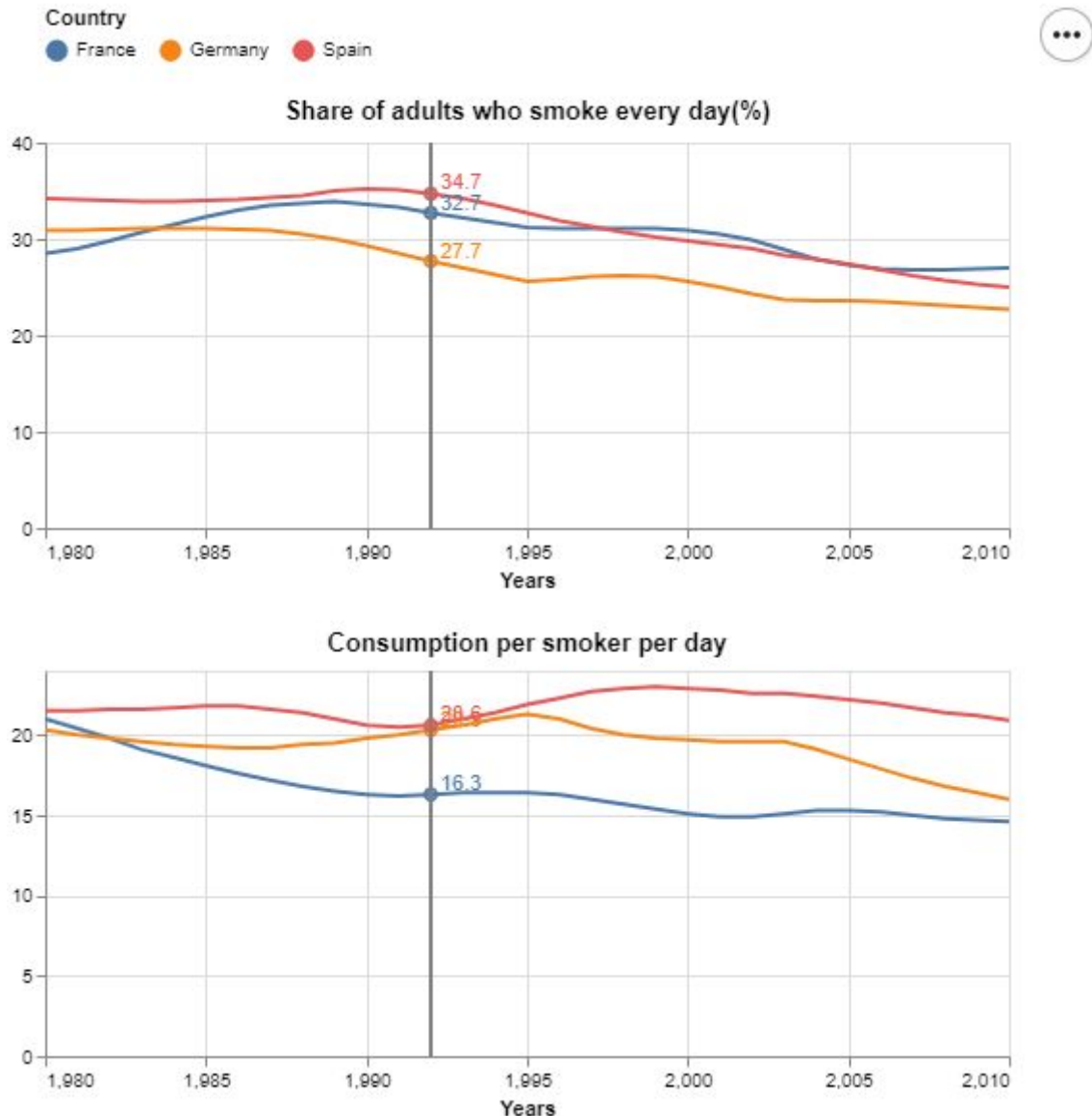
The first chart is designed to demonstrate the seriousness of the issue. It doesn't exaggerate or diminish the problem, it shows the exact percentage of smokers which lets the viewer understand the real situation. The second chart is a logical consequence of the previous chart; once the user had seen the share of smokers, they may question themselves about the amount of tobacco the smokers consume, and that is precisely what the second chart shows. To summarise, the first chart shows the scale of the issue, the second chart zoomes into the problem on the individual level.

Smoking Deaths

The first question that we want to answer is: Is Tobacco a deadly threat? To answer this question, we may want to explore different countries during a particular period of time in terms of smoking deaths. The country can be selected in the drop down list and the years are chosen by using the slider as below.

Select a country:

Germany ▾

Select a period of time

1994   2004
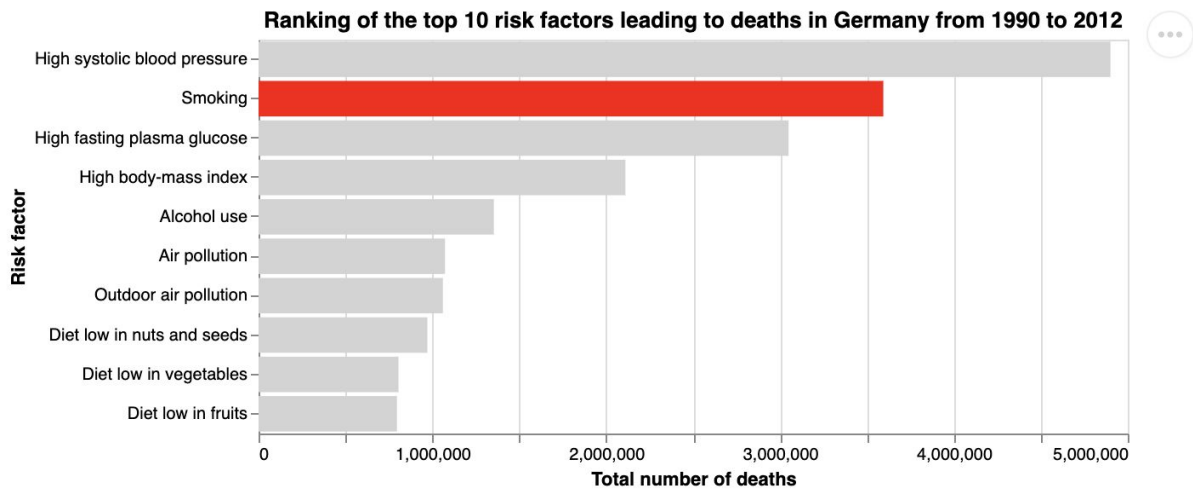
1990   2012

We use a stacked bar chart with three different age ranges to visualize the smoking deaths by age. The number of smoking deaths aged 15 to 49 is in red, 50 to 69 is in orange, and above 70 is in light orange. The x-axis is showing the years and the y-axis is representing the number of people died because of smoking. The tooltip showing the number of deaths for a specific group of people will be shown when users hover over the bars. We intentionally chose this color scheme to make the bars look like cigarettes.

Smoking deaths by age in Germany
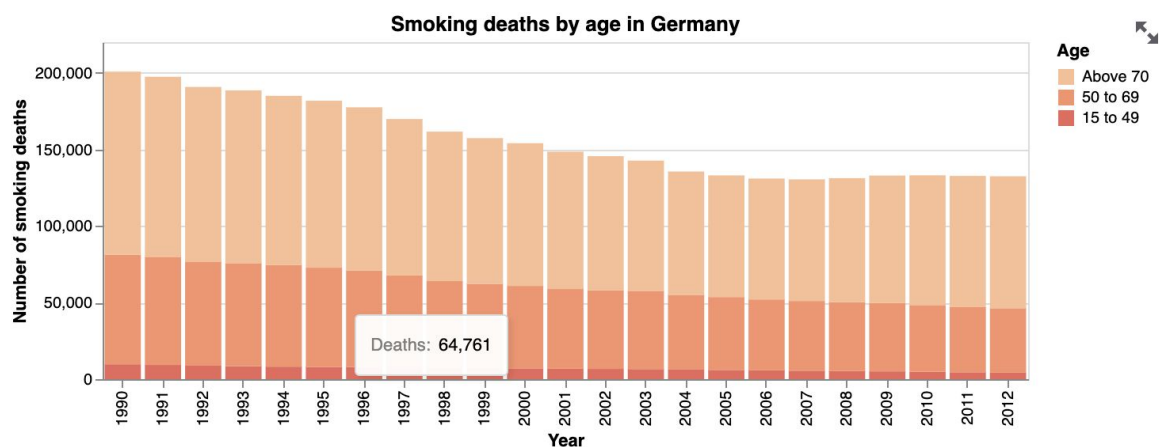
The second question we want to answer is: Is smoking the most dangerous factor leading to deaths? To answer this question, we are putting smoking next to other critical factors to be able to compare them. On the left, the y-axis is showing the top 10 factors leading to deaths and the x-axis is the total number of deaths in the chosen country over the chosen period of time. The tooltip showing the total number of deaths for a specific factor will be shown when users hover over the bars.

Ranking of the top 10 risk factors leading to deaths in Germany from 1990 to 2012

Risk factor	
High systolic blood pressure	
Smoking	
High fasting plasma glucose	
High body-mass index	
Alcohol use	
Air pollution	
Outdoor air pollution	
Diet low in nuts and seeds	
Diet low in vegetables	
Diet low in fruits	

Total number of deaths: 0, 1,000,000, 2,000,000, 3,000,000, 4,000,000, 5,000,000

Tobacco Control

The following analysis is based on the evaluation made by the World Health Organization (WHO) to country policies against Tobacco. WHO classified these policies in 6 categories detailed in the following table:

Policy	Description
Monitor	Monitor tobacco use and prevention policies. There are repeated significant surveys on Tobacco issues for each country.
Protect from Tobacco Smoke	Evaluates how countries implement Smoke-free environments in order to prevent both direct and indirect smoking risks.
Offer help to quit Tobacco use	Evaluates the availability and non-availability of particular tobacco cessation aids and assistance to help tobacco users quit
Warn about the dangers of Tobacco	Countries' legislation is evaluated to determine if health warning with specific criteria is provided to the public
Enforce bans on Tobacco advertising	Countries' legislation is evaluated to determine if direct and indirect advertising promotion or sponsorship for Tobacco is banned or regulated.
Raise taxes on Tobacco	Measures the % of taxes of the retail price of Tobacco for each country

A score from 1 to 5 is assigned depending on the intensity of a country to deal with Tobacco issues being 1 the worst and 5 the best. The overall analysis is based on a policy selected by the user:

Select control measure:

Monitor

Monitor

Protect from tobacco smoke

Offer help to quit tobacco use

Warn about the dangers of tobacco

Enforce bans on tobacco advertising

Raise taxes on tobacco

After selecting the policy or control measure, the user can navigate and discover insights throughout three different visualizations:

1. Implementation of the policy around the world

The main objective of the visualization is to allow the user to identify the intensity of a specific control policy for each country. This visualization consists of a map which shows the implementation score of the selected policy in each country filtered by a year the user can select from the select box on the top.

Each country is encoded by a spatial region within the map which allows to have a global view of the problem as well as patterns among countries from the same continent. The implementation score value is encoded by color saturation being the best implementation the strongest color and the worst the softest color.

In addition, The user can also identify the description of each color in the legend located in the bottom part of the visualization. the user can also look for the detail of a region by hovering over the desired space. It will display a tooltip with the description of the country.

Implementation of the policy "Monitor" around the world

Monitor

| 13%	27%	27%	32%
1.Very Bad	2.Bad	3.Medium	4.Good
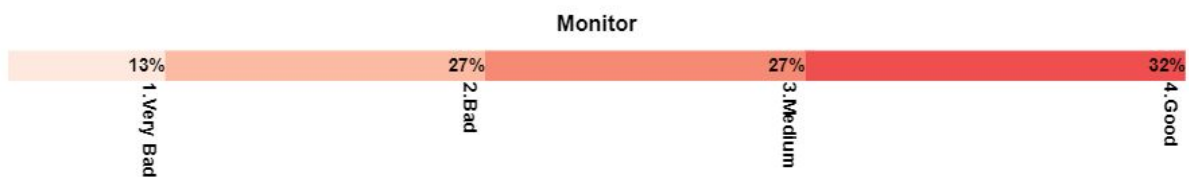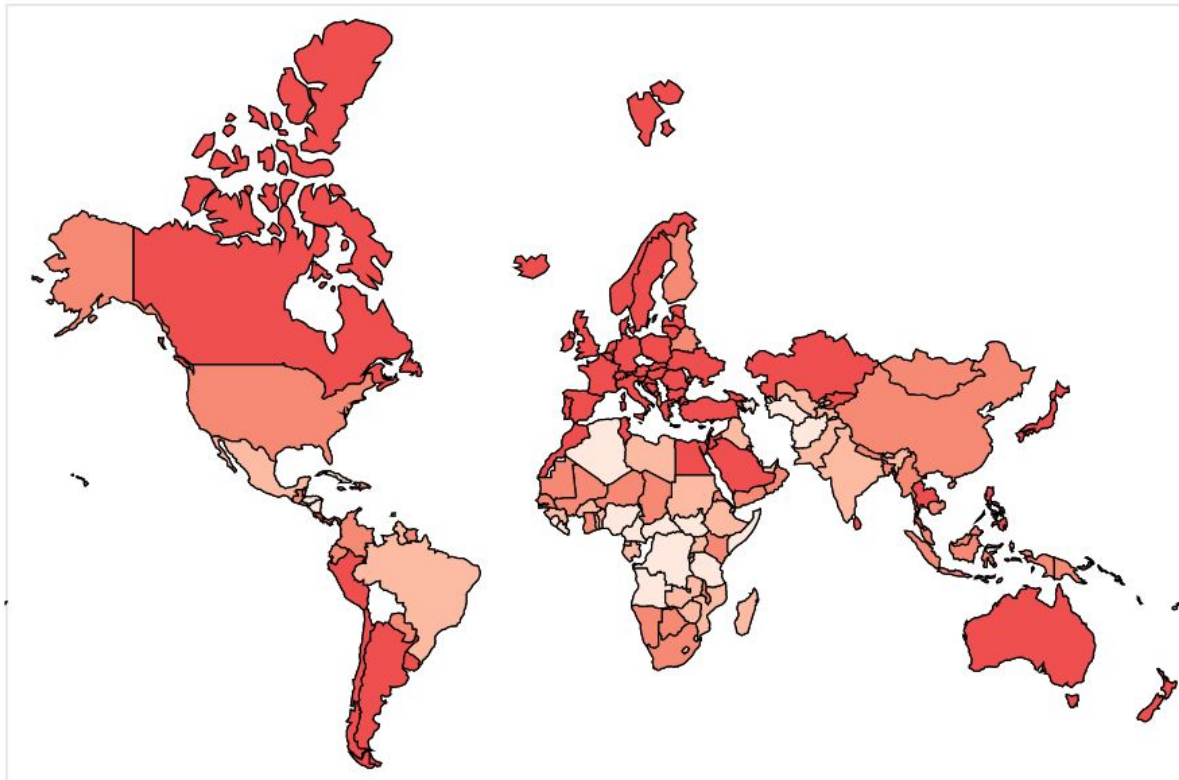
Another interesting feature is the information that the legend displays. The legend is a stacked barchart whose size encodes the percentage of countries corresponding to the respective implementation score for the policy. The user can identify whether a selected control policy is well implemented or not.

Protect from tobacco smoke

3%	58%	25%	7%	8%
1.Very Bad	2.Bad	3.Medium	4.Good	5.Perfect
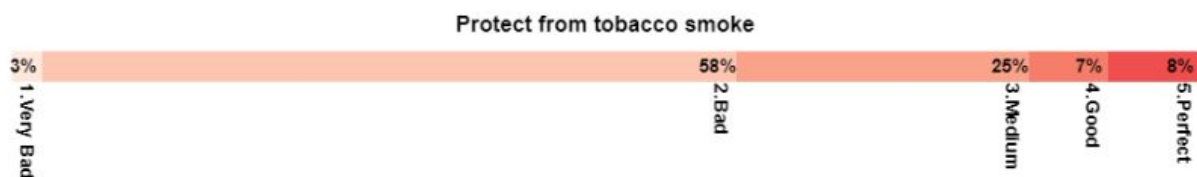
2. Evolution of the policy per country

This visualization aims to evaluate and compare the periodic evolution of the selected policy among different countries of our interest. The user can select many countries in the select box on the top. The chart will automatically increase its vertical size when having more countries to compare. In this case the countries are encoded by rows and the years are encoded by columns. The implementation score value of the selected policy is encoded by both the size of the horizontal bar and color saturation. Finally, the user can hover over the bars in order to check the tooltip information.

As in the case of the map, the user can also visualize the legend on the bottom part of the chart in addition to the percentage of countries that each score represents.
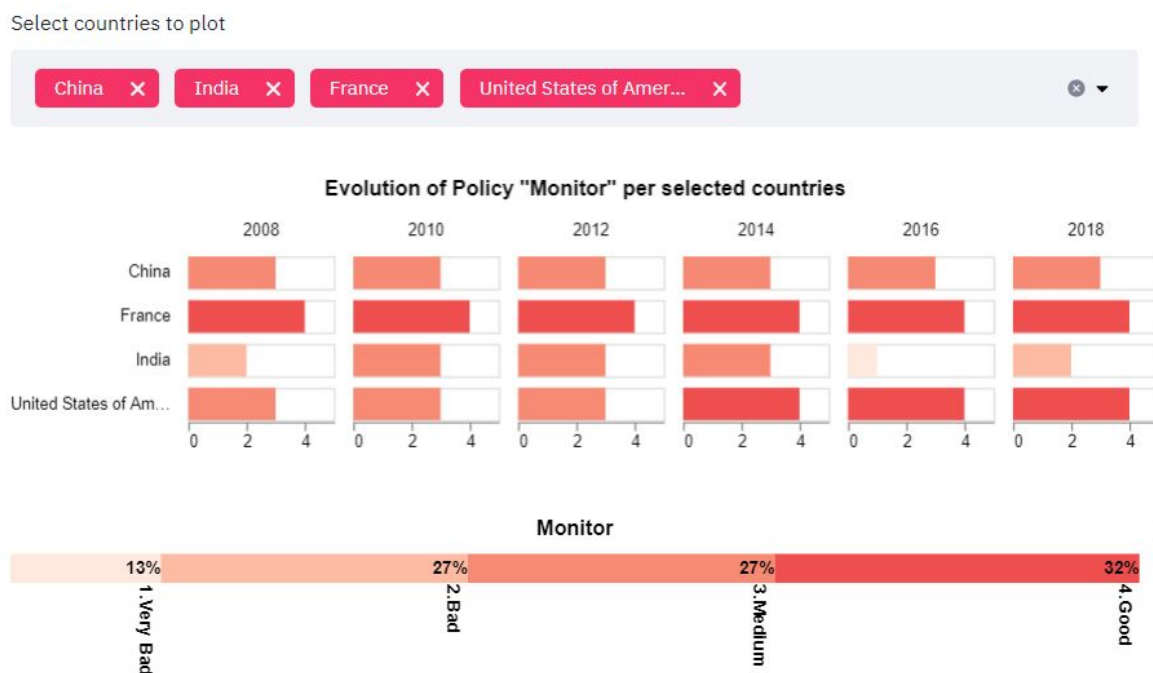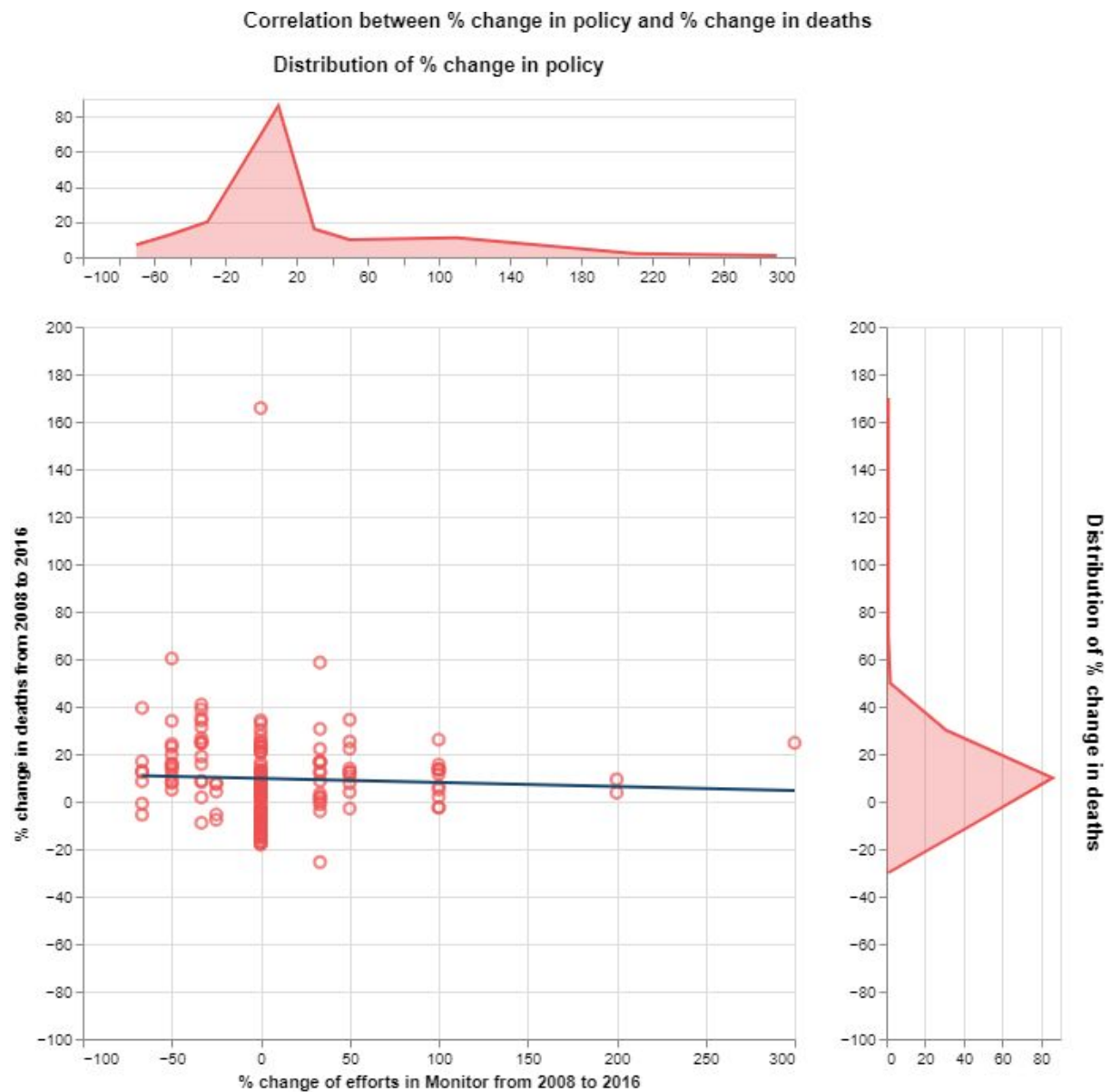

3. Correlation between % change in policy and % change in deaths by smoking

Countries have implemented different control policies against Tobacco which have been measured by WHO from 2008 until 2018. During this period, some countries have strengthened their policies; however, we don't know the real impact of them. The last visualization tries to identify correlations between the change in policy and the change in deaths by smoking from 2008 to 2016. The definitions of each policy are updated automatically each time we select the control policy:

$$\% \ change \ in \ Monitor \ = \ \frac{Monitor \ in \ 2016}{Monitor \ in \ 2008}$$

$$\% \ change \ in \ Deaths \ by \ Smoking \ = \ \frac{Deaths \ by \ Smoking \ in \ 2016}{Deaths \ by \ Smoking \ in \ 2008}$$

The visualization consists of 3 charts which tends to provide an analytical view of the correlation of these variables. The main chart is a scatterplot whose points encode the countries. The axis X represents the % change in a respective policy and the axis Y represents the % change in deaths by smoking. Likewise, the histogram on the top also represents the distribution of % change in a respective policy and it's located in parallel to the axis X for better understanding. The same occurs with the histogram on the right which represents the distribution of the % in deaths by smoking and it's parallel to the axis Y. Another interesting feature is the slope line which represents the regression line between the two variables. It allows the user to have an idea of the trend of the correlation.

Correlation between % change in policy and % change in deaths

Distribution of % change in policy

% change in deaths from 2008 to 2016

% change of efforts in Monitor from 2008 to 2016

Distribution of % change in deaths

The user can also select brush the histograms in order to filter the points and evaluate the slope of the regression in more detail (with groups that increased more or less in control policies, for example). As a result, the country points will be filtered in the scatterplot and the slope line will be updated.
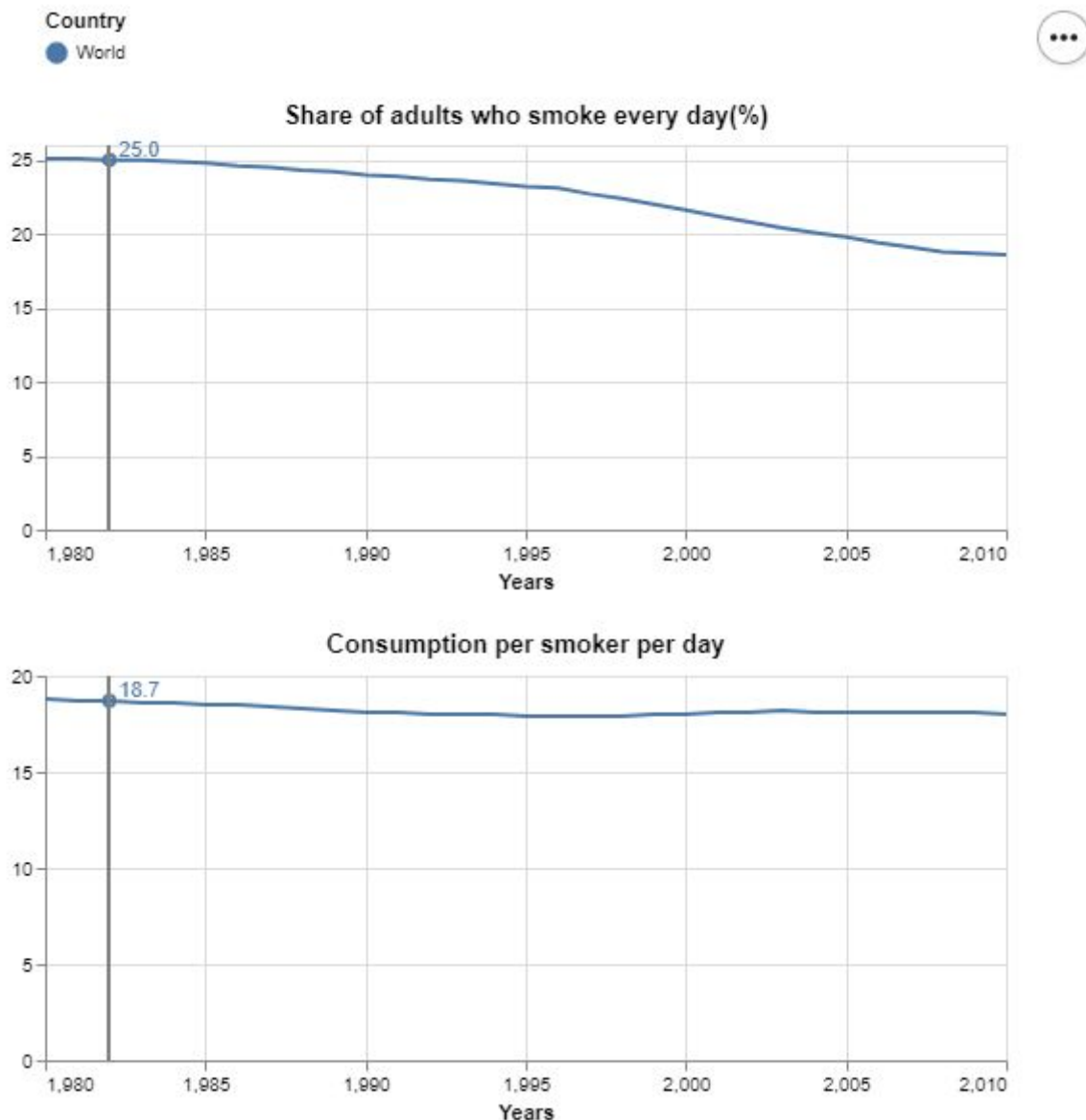
Correlation between % change in policy and % change in deaths
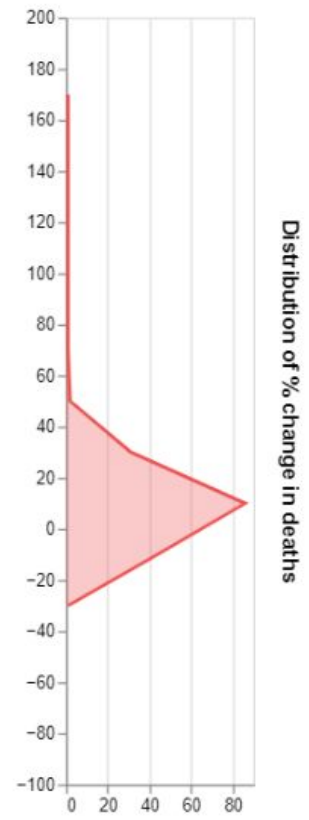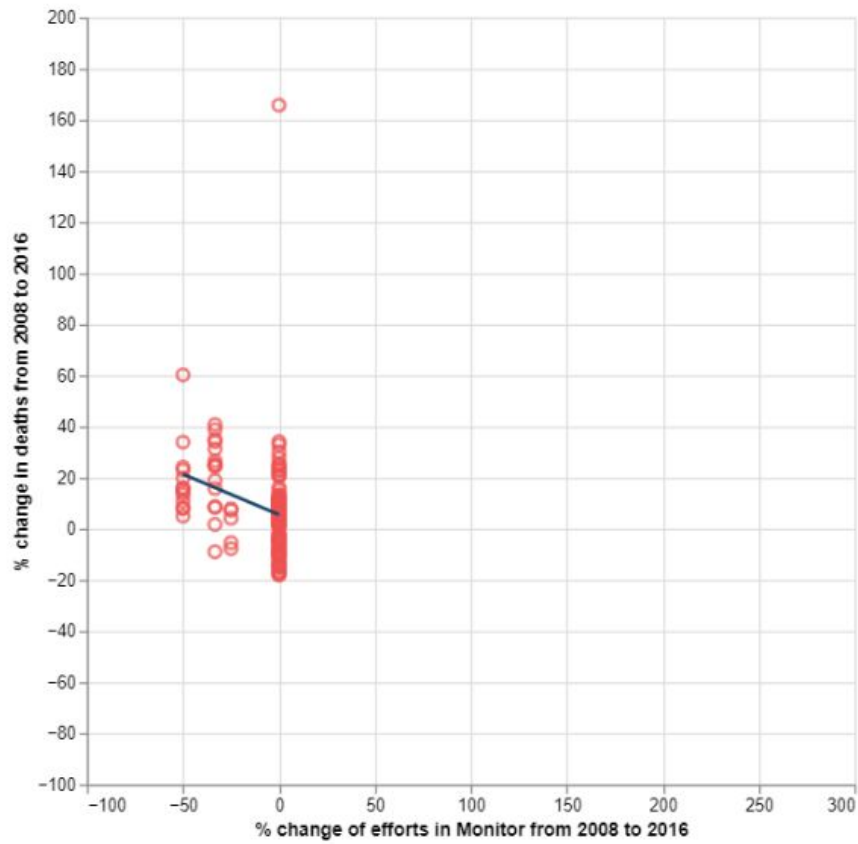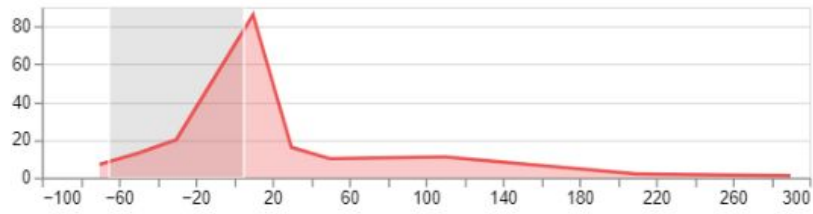
4. Interesting Insights

Tobacco Consumption

Country
● World

Share of adults who smoke every day(%)

Consumption per smoker per day
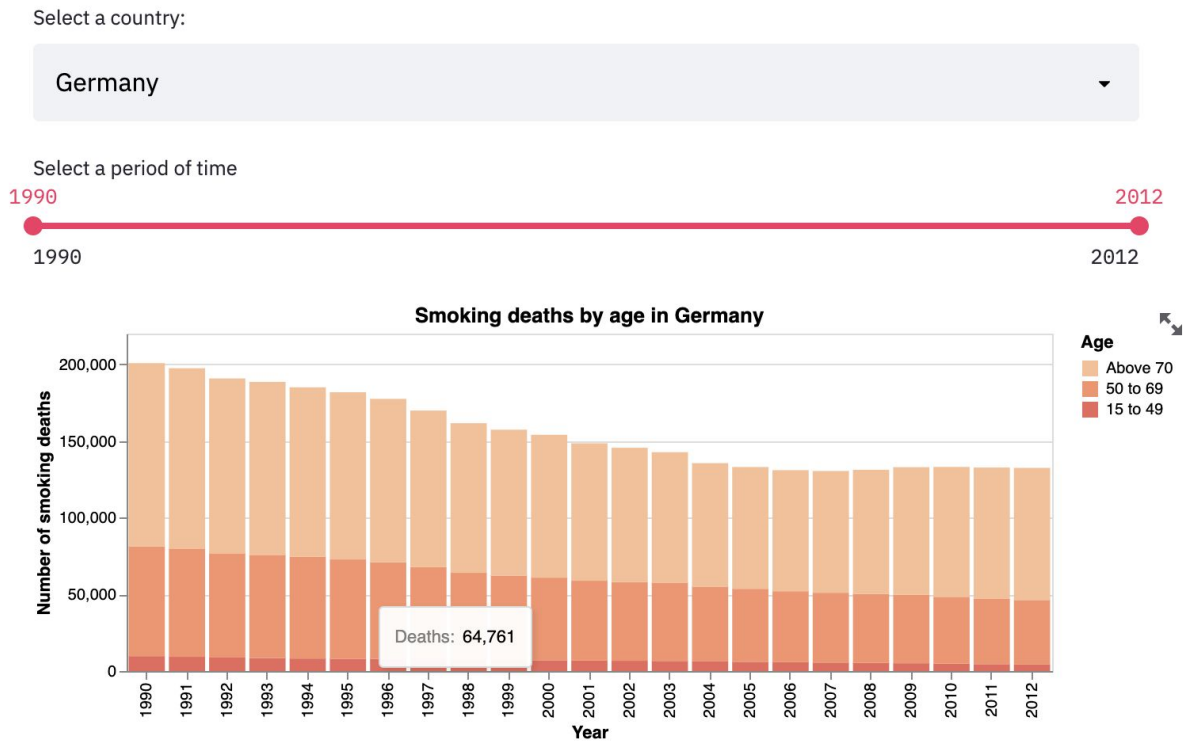

It is intuitive that, overall, in the world, smoking is becoming less popular, and the first chart proves that statement. However, the value of the y-axis of the second chart (Consumption per smoker per day) doesn't change much across the whole timeline; what does it suggest? One can make a hypothesis that the number of new smokers is decreasing every year but most of those who smoke don't change their smoking habits. In fact, the first part of the hypothesis is not what we should focus on; there can be several factors which affect the share of every-day-smokers: improvement in teenage smoking prevention or powerful anti-tobacco policies; the interesting insight is that smokers do not change their habit of smoking even in 30 years. Of course, it is a generalising statement, certainly, there are people who reduce their

consumption of cigarettes, but they do not affect the big picture. The big picture suggests that tobacco consumers either stop smoking abruptly (death or just stopped smoking) which means they do not count as smokers anymore, or they keep smoking with the same frequency. It is a sad insight but it is important to realise.

Smoking Deaths

Select a country:

Germany ▾

Select a period of time

1990   2012

1990   2012

Smoking deaths by age in Germany

Deaths: 64,761

Age
- Above 70
- 50 to 69
- 15 to 49

It can be clearly seen that overall the number of people who died because of smoking in Germany is decreasing over this period of time. When we look at the breakdown of deaths from smoking by age we see that older populations are affected the most. Here we see that death rates from smoking are much higher in people older than 70 years old, followed by those aged 50 to 69. Death rates for smoking people aged 15 to 59 are low.

Let's have a bigger picture by showing the data for the whole world. Overall, the number of smoking deaths is slightly increasing over this period of time. However, in this chart, we have not taken into account the fact that the population of the whole world is changing over the years. In 2012, Smoking was responsible for almost 7 million deaths. So imagine on average, everyday there are 20,000 people who die because of smoking all over the world.

World ▾

Select a period of time
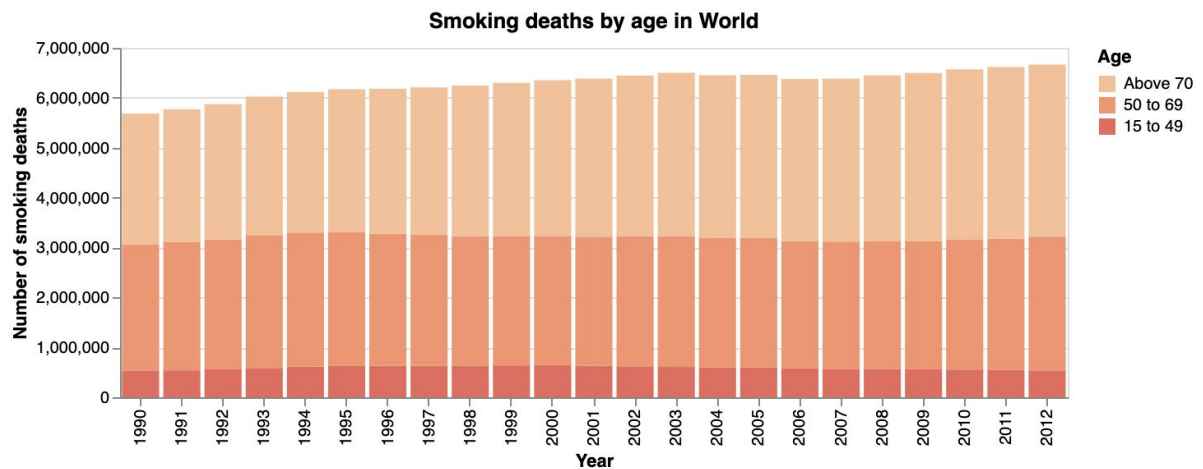
1990   2012

1990   2012

Smoking deaths by age in World
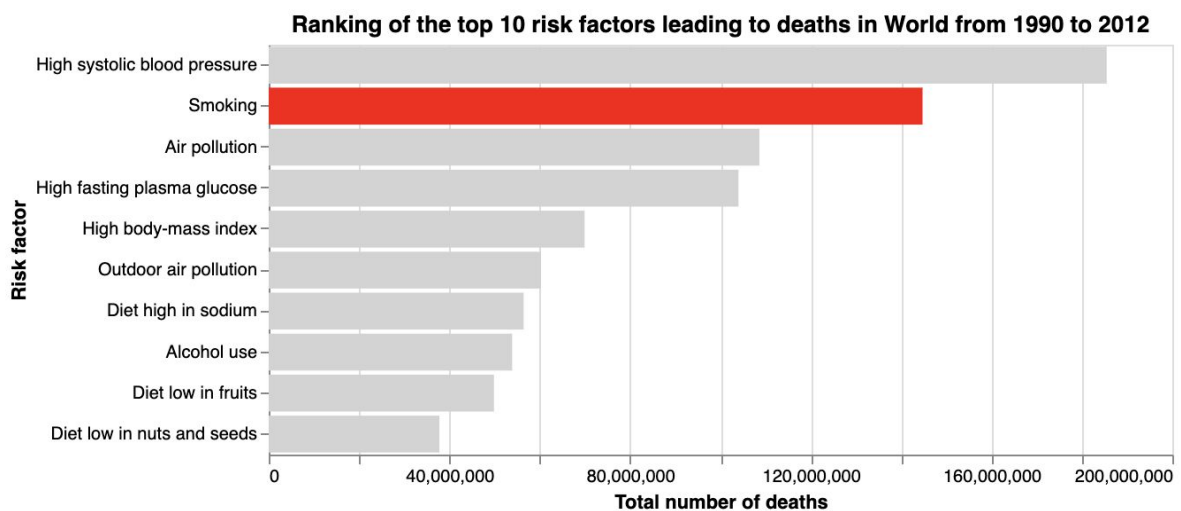

In the bar chart below, we can see that smoking is one of the world's largest health problems. More specifically, it ranks second in the world during the period of time from 1990 to 2012. The only one factor, which ranks higher than smoking in this list, is high systolic blood pressure.

Ranking of the top 10 risk factors leading to deaths in World from 1990 to 2012

One more interesting insight that we found is in France. As shown in the chart below, during the period of time from 1990 to 2012, smoking ranks first on the risk factor list, which means it is the most dangerous factor leading to death in France.

Ranking of the top 10 risk factors leading to deaths in France from 1990 to 2012
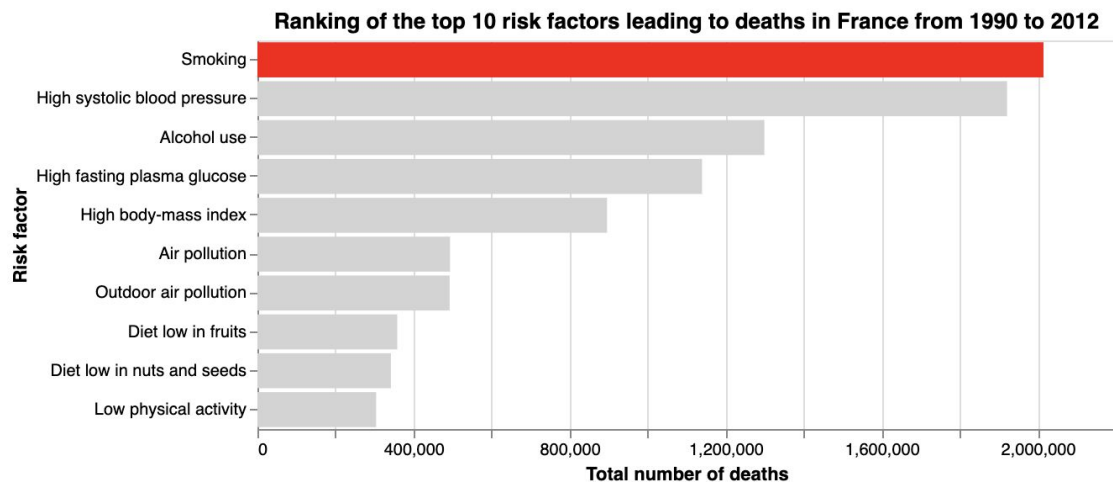

Tobacco Control

First of all, the map allows us to identify a joint policy of european countries for **monitoring** Tobacco in 2018. In fact, the European Union [3.1] has a well developed strategy of legislations, recommendation and information campaigns to prevent Tobacco consumption which is reflected in this visualization. On the other hand, countries in Africa don't make efforts on monitoring in the recent survey. Overall, only 40% of the countries efficiently monitors Tobacco issue as can be noticed in the legend of the bottom

Implementation of the policy "Monitor" around the world

Monitor

| 14%	30%	18%	38%
1.Very Bad	2.Bad	3.Medium	4.Good
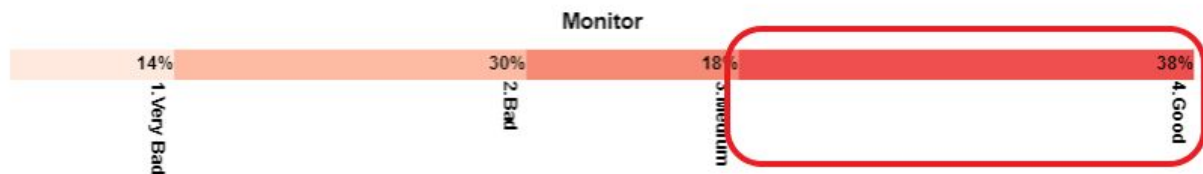
Other interesting insights can be derived from the policy **"Warn about the dangers of Tobacco"**. In 2008, only 16% of the countries really warned the population about the risks of Smoking Tobacco.
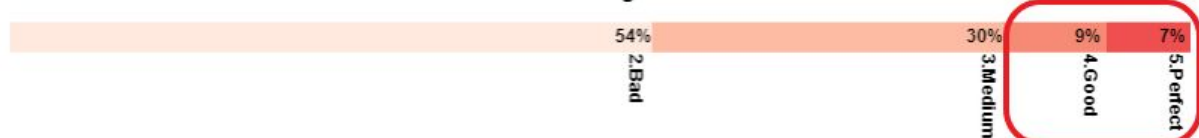
Implementation of the policy "Warn about the dangers of tobacco" around the world

Warn about the dangers of tobacco

| 54% | 30% | 9% | 7% |
| 2.Bad | 3.Medium | 4.Good | 5.Perfect |
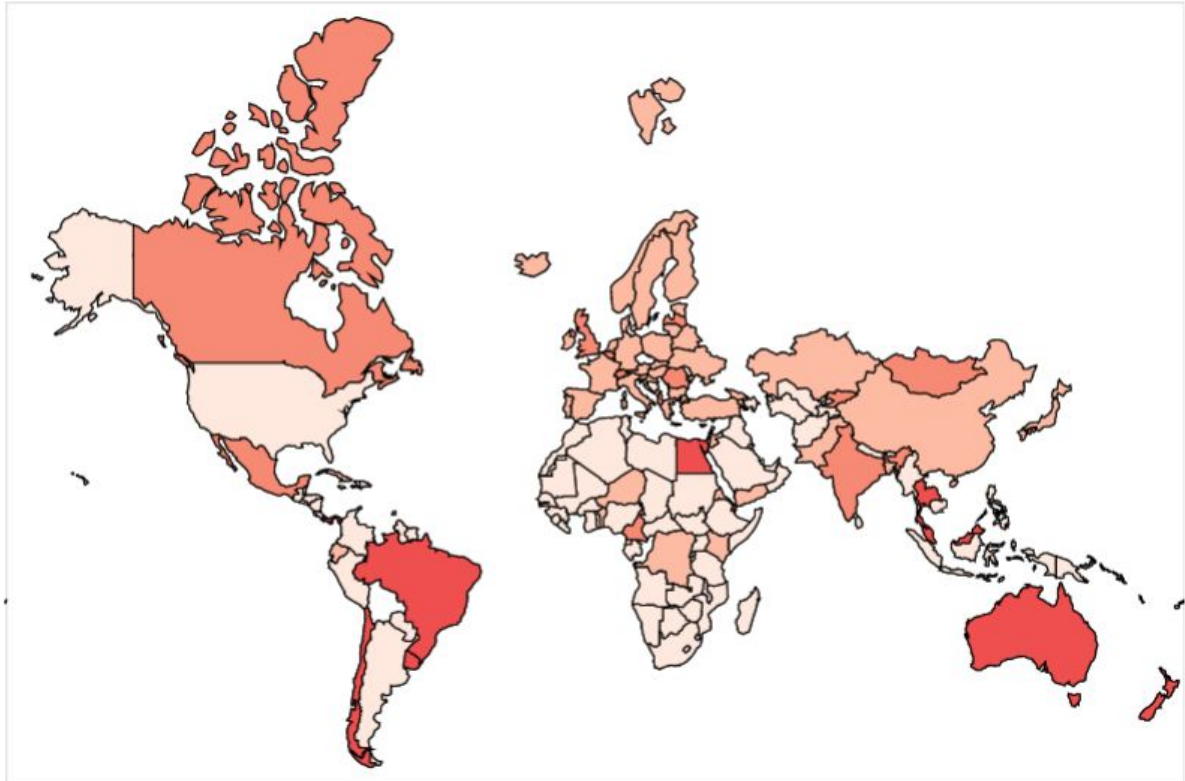
Since WHO started to deep into the problem as an epidemic, the world took more consciousness about this issue. As a consequence, the percentage of countries implementing this policy **grew to 63%** in 2018, the recent year of analysis. In other words, almost 50% of the world took Smoking Tobacco as a serious problem.

Implementation of the policy "Warn about the dangers of tobacco" around the world

Warn about the dangers of tobacco

22%	16%	15%	47%
2.Bad	3.Medium	4.Good	5.Perfect
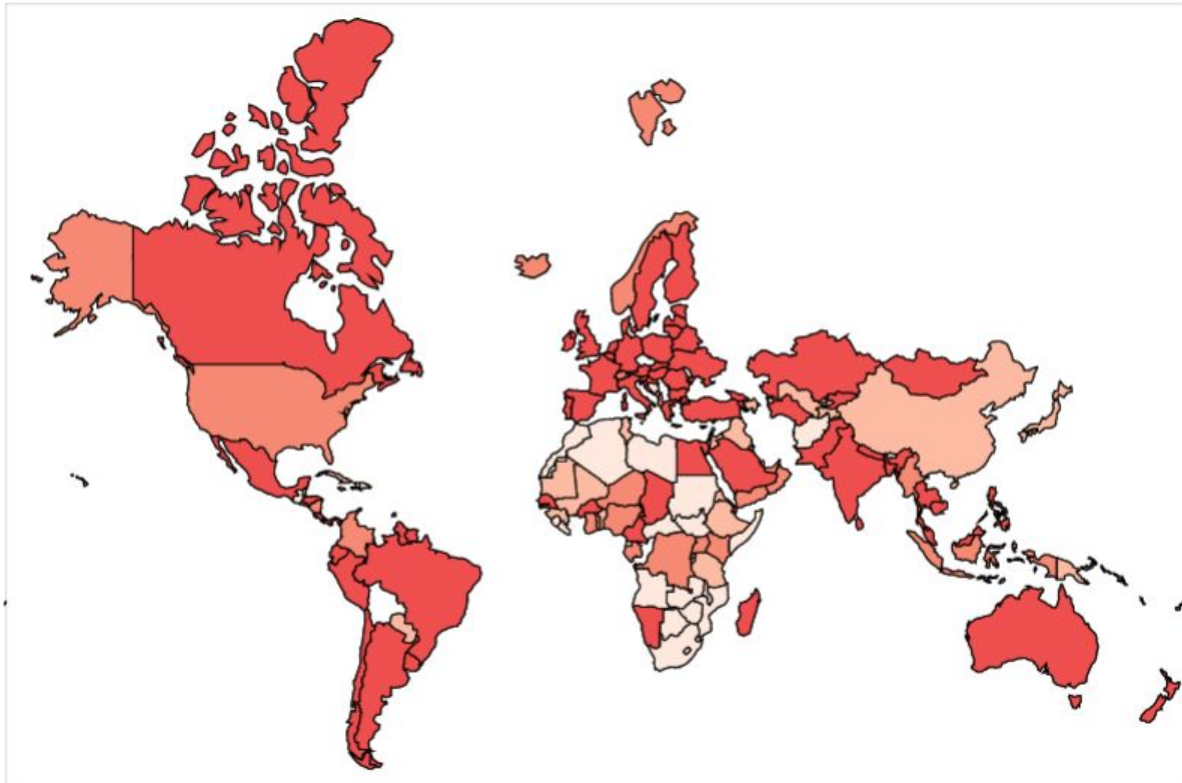
Regarding the evolution of the policy **"Protect from tobacco smoke"** per period, we have selected some countries of our interest and compare the score of their implementation. Brazil can be considered an example in the fight against Tobacco. From 2012, they have implemented high policies to protect free smoke environments. Likewise, India keeps their good measures in this policy. On the other hand, France registers a low strategy to propose laws which protect people of being affected by passive Smoking.

China ✕   India ✕   France ✕   Brazil ✕   United States of Amer... ✕   Germany ✕   ⊗ ▾

Evolution of Policy "Protect from tobacco smoke" per selected countries

	2008	2010	2012	2014	2016	2018
Brazil						
China						
France						
Germany						
India						
United States of Am...						

Protect from tobacco smoke

5%	27%	22%	14%	32%
1.Very Bad	2.Bad	3.Medium	4.Good	5.Perfect
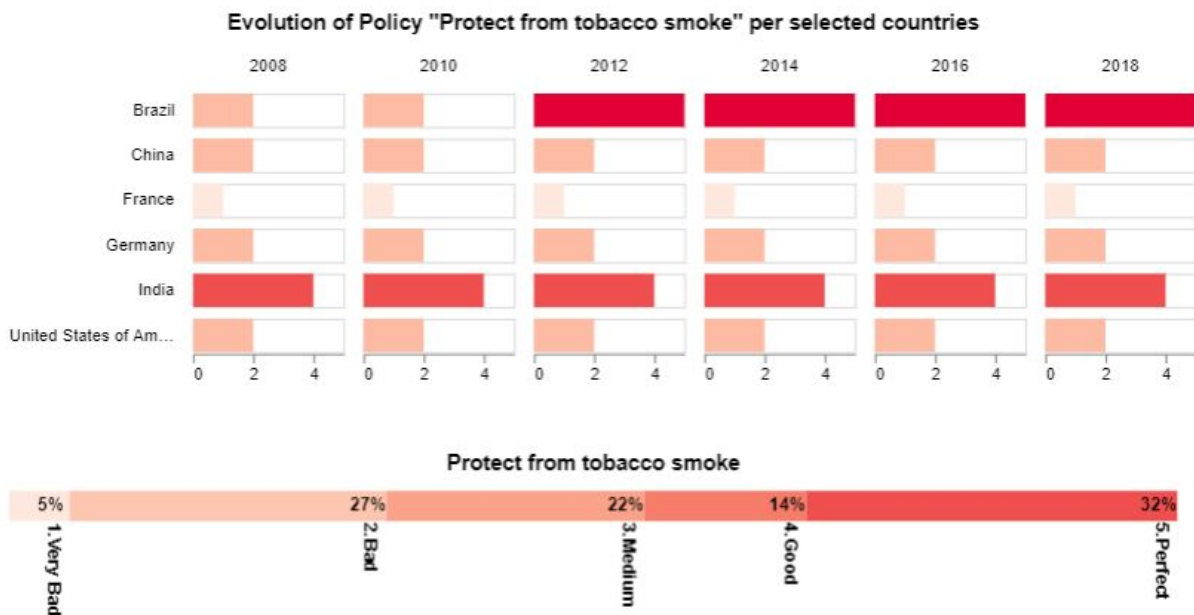
While looking at other policies such as **"Raise taxes on tobacco",** we can see that most of the selected countries improved and fixed high taxes on tobacco over time with the exception of the United States. Brazil keeps a perfect implementation of these measures contrary to the US.

China ✕   India ✕   France ✕   Brazil ✕   United States of Amer... ✕   Germany ✕   ⊗ ▾

Evolution of Policy "Raise taxes on tobacco" per selected countries

	2008	2010	2012	2014	2016	2018
Brazil						
China						
France						
Germany						
India						
United States of Am...						
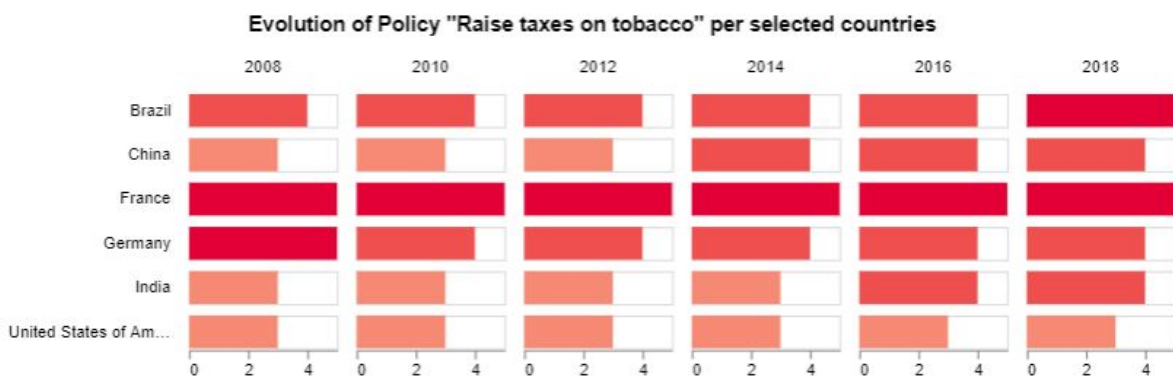
In the next chart, we can also see that most of the countries focus on **"enforcing bans on tobacco advertising".** One more time, Brazil is the example with the best implementation in the majority of the policies until 2018. On the other hand, the
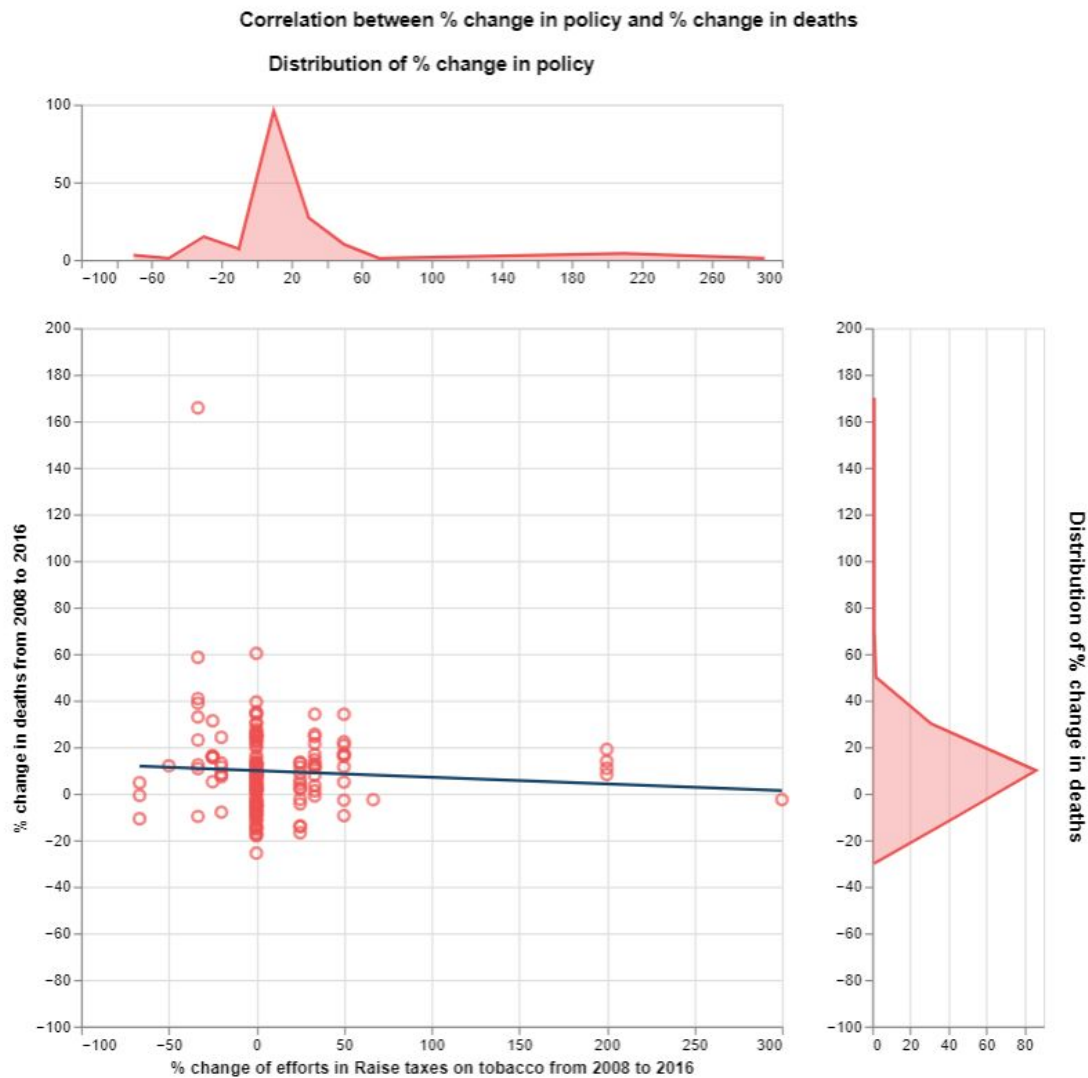
United States did not address its efforts to ban Tobacco advertising. The US can be considered one of the countries with less commitment to fight against this issue.
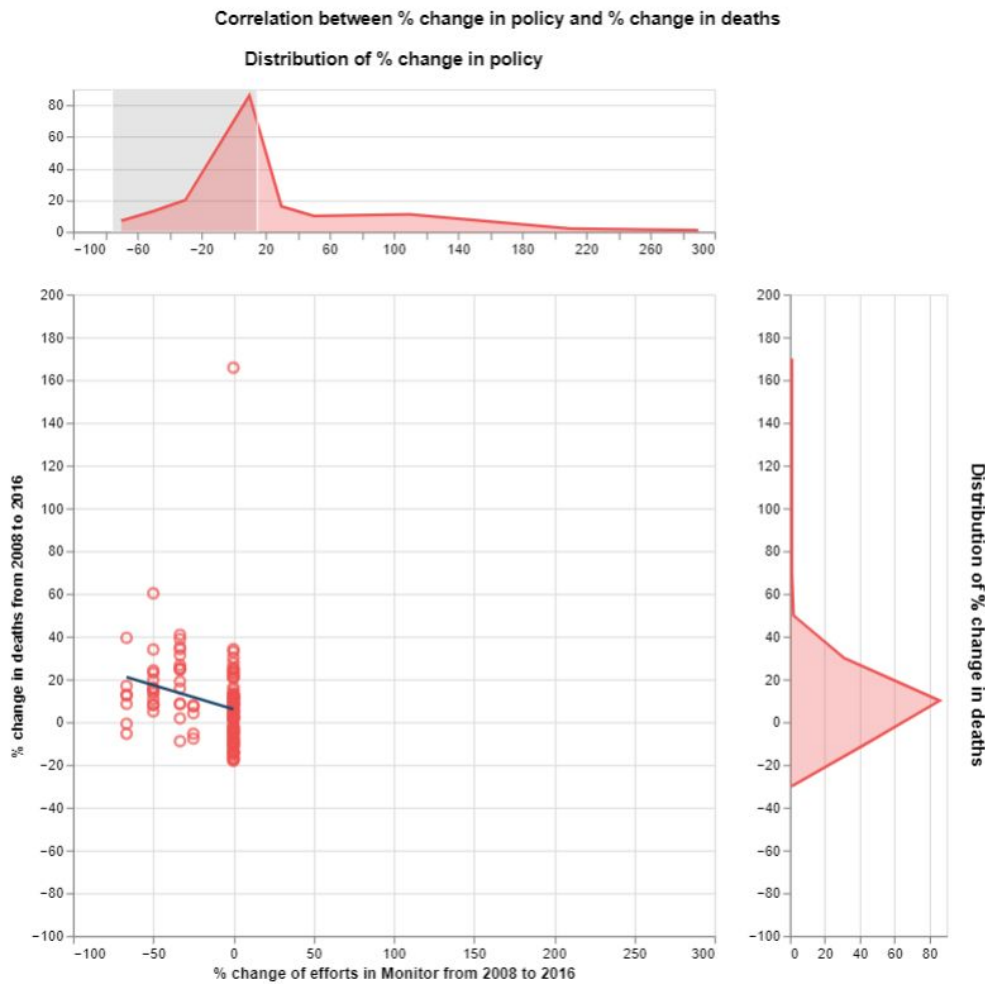

From the last visualization, the correlations allow us to identify some trends between the variables. The most interesting insight is the correlation of the percentage of change in the policy **"Raise taxes on tobacco"** and the percentage of change in **"deaths by smoking".** The **regression line** clearly defines a trend between these two variables which can be described as: countries which raised taxes on tobacco have on average less increase in the number of deaths. As a consequence, this policy would be ideal to fight with the Tobacco Problem. In fact, there are studies which suggest this relationship [3.2]

Correlation between % change in policy and % change in deaths

Distribution of % change in policy

% change in deaths from 2008 to 2016

% change of efforts in Raise taxes on tobacco from 2008 to 2016

Distribution of % change in deaths

Finally, a stronger effect can be observed when brushing the histograms and analysing the countries which **reduced taxes on Tobacco.** These countries increased the number of deaths by smoking thereby representing a major issue in the country's strategies. While correlation doesn't mean a direct cause, this policy seems to be important in the decision-making process to fight against Tobacco.

Correlation between % change in policy and % change in deaths

Data Sources

- Tobacco Epidemic: https://www.who.int/news-room/fact-sheets/detail/tobacco#:~:text=Tobacco%20kills%20more%20than%208,-%20and%20middle-income%20countries.
- Tobacco Consumption: http://ghdx.healthdata.org/record/ihme-data/global-smoking-prevalence-and-cigarette-consumption-1980-2012
- Smoking Deaths: http://ghdx.healthdata.org/gbd-results-tool
- Tobacco Control: https://apps.who.int/gho/data/node.main.Tobacco?lang=en

References

[3.1] https://ec.europa.eu/health/tobacco/overview_en
[3.2] https://tobaccocontrol.bmj.com/content/11/suppl_1/i62.full