Notes:
- We uploaded the python code (of visualizations) to our repository, under the folder named 'Python Visualizations' in two separate notebook files.
- We uploaded the output visualizations of the code separately under the 'Python Visualizations'/'Outputs' folder.

GitHub repository for the project:

https://github.com/ihagverdi/Database-Systems-Project

Group 2:

Hagverdi Ibrahimli - 30014

Raman Afravi - 30061

Amil Kazımoğlu - 27891

Ayçelen Kaptan - 28826

Kaan Özyer - 26837

Project title and description:

Our project title is "Smoking kills" and the main idea behind the project is to examine the relation between smoking and health. The data we shall utilize in this database application covers the following topics: Affordability of cigarettes, consumption per smoker per day, cigarette sales per day, share of cancer deaths due to tobacco, share of deaths due to smoking, share of lung cancer deaths. The real world problems which this database application addresses narrow down to uncovering the negative relationship between the smoking rate of the individuals and their health.

**VISUALIZATIONS**

Graph 1 (Scatter plot):

We calculated the average number of deaths due to tobacco for all countries in our dataset and applied the same procedure to the cigarette consumption rate of the countries. Then, we used a scatter plot to visualize the relationship between these two variables: Average cigarette consumption and average tobacco deaths. It is important to note that deaths due to tobacco also include the deaths due to smoking.
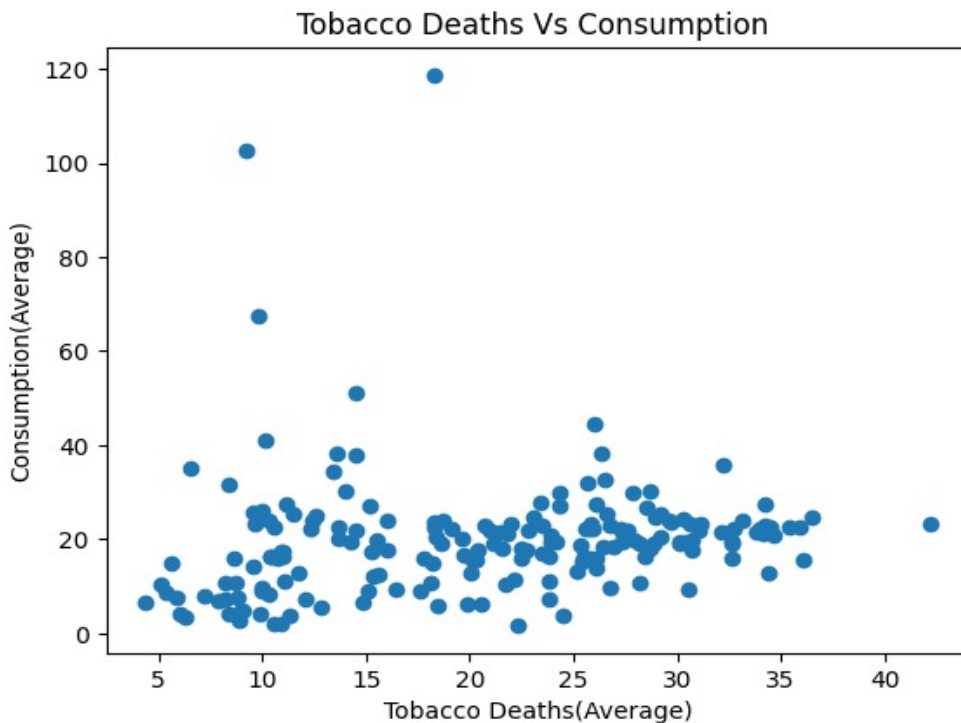
From the graph above, we can see that the average number of deaths due to tobacco usage (which includes smoking as well) is positively correlated with average consumption rate for the world countries. This finding suggests that higher the

consumption of tobacco, the more likely it is to have a higher death rate from tobacco which in turn uncovers the negative effect of tobacco on the population's health.

Graph 2 (Column chart):
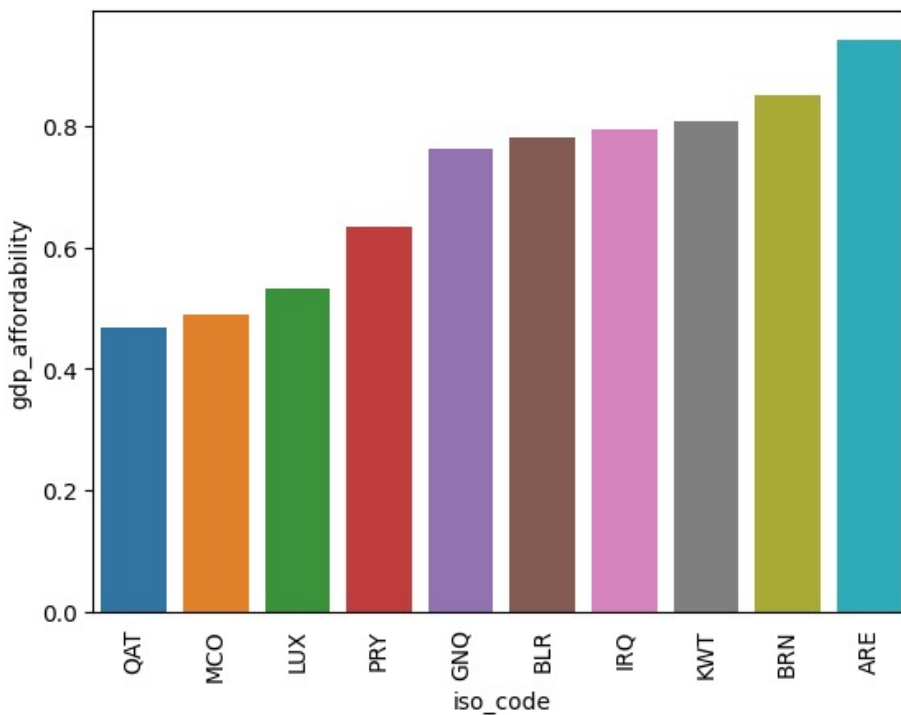
We calculated the share of the average GDP per capita required to purchase 100 packs of the most-sold brand of cigarettes for all countries in our dataset and graphed the top 10 countries in a column chart. It is important to note that, the higher the gdp_affordability the more the share of the GDP per capita required to purchase the cigarettes. This graph can be used to make a relationship between how much cigarettes are consumed versus how much people can afford.

Graph 3 (Area chart):
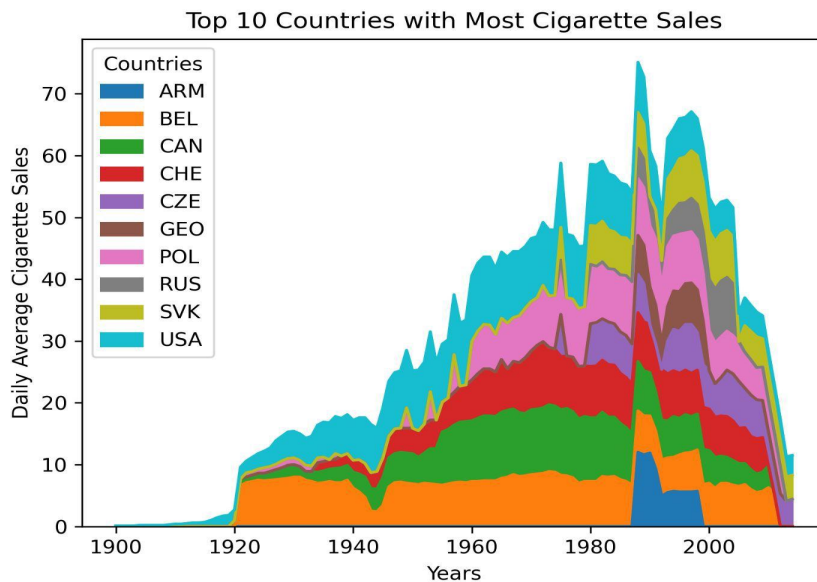
This chart shows the top 10 countries with the highest number of daily cigarette sales. It describes the sales of cigarettes in world countries between the years 1900-2020. From this chart we can see that the sales of cigarettes peaked around the 1990's also around the time when cigarette consumption was at its highest. From the Cigarette Consumption line chart we can see that both cigarette sales and cigarette consumption sharply increases approximately around the same time.

Graph 4 (Bar chart):

Top 10 Countries With the Most Lung Cancer Deaths Due to Smoking
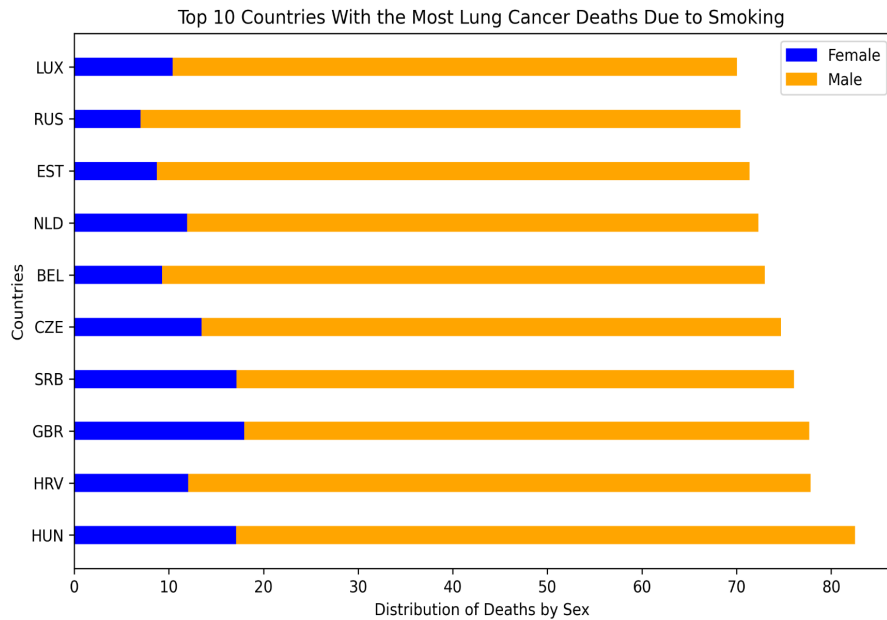
This is a bar chart of Countries with the highest rate of lung cancer deaths and their distribution between sexes. From this graph we can argue that there is a correlation between sexes, their cigarette consumption and their death rates of lung cancer. Males have a higher rate of death than females so we can conclude that smoking is more popular between males. We can also see from the Cigarette Consumption line chart that as consumption increases lung cancer deaths also increases.

Graph 5 (Line chart):

Cigarette Consumption
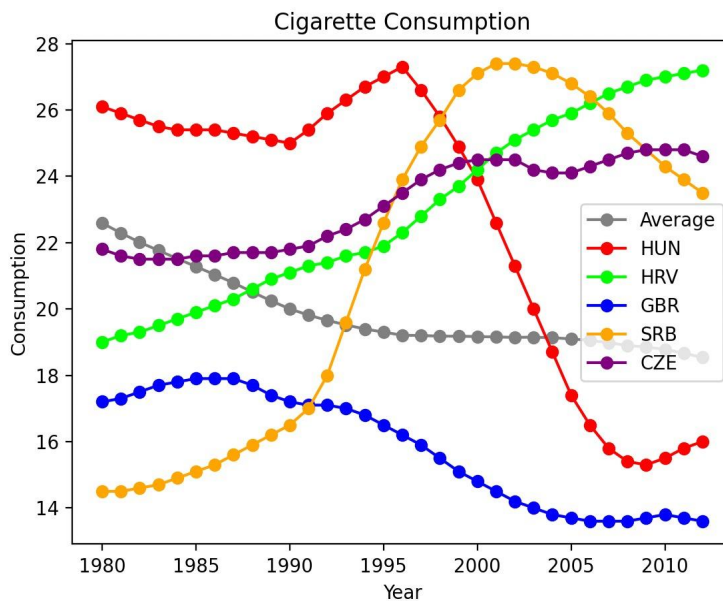
Then we are using the countries that we found previously that are known to have the most deaths resulting from lung cancer to create a line-chart, comparing their respective cigarette consumption throughout the years. We can conclude from both charts that high cigarette consumption alone is not the leading cause of death, and other factors should also be considered.

Graph 6 (Pie chart):

Top 10 Countries With Highest Cigarette Consumption
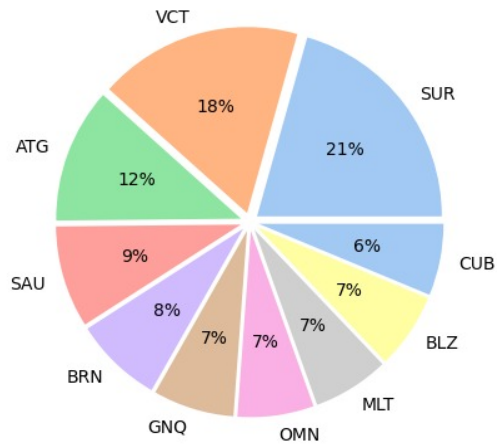
Lastly we retrieved the countries with the highest cigarette consumption. The pie chart shows the 10 countries with the highest cigarette consumption. This pie chart is utilizing all the dates we have in the dataset and using them to calculate the top 10 countries which consume most cigarettes. We wanted to clarify if consumption alone does in fact result in lung cancer death, it can be observed that the top cigarette consuming countries are different from the countries with the highest number of deaths due to lung cancer.