

Project 1: World Development Statistics

**Tobacco Usage vs Life Expectancy
in 2020**

Sophia Joseph





Problem Statement

Smoking is a very prevalent problem in various countries and regions across the world and can be the source of serious health problems and can ultimately lead to death in many cases. The objective of this project is to find whether countries with a low life expectancy also have a high rate of smoking in the year 2020.



Steps to tackle the problem:

- 1) Import datasets for usage
 - Population.csv - Population through countries
 - Life_expectancy.csv - Life Expectancy through countries
 - Share-of-adults-who-smoke.csv - Percentage of adults who use tobacco
- 2) Data cleaning
- 3) Exploratory Data Analysis
- 4) Data visualization



Data Cleaning

- 1) Check for and remove any unnecessary missing values
- 2) Change dataframe column names
- 3) Fix incorrect data types
- 4) Apply convert method to datasets
- 5) Merge datasets and save them



Exploratory Data Analysis

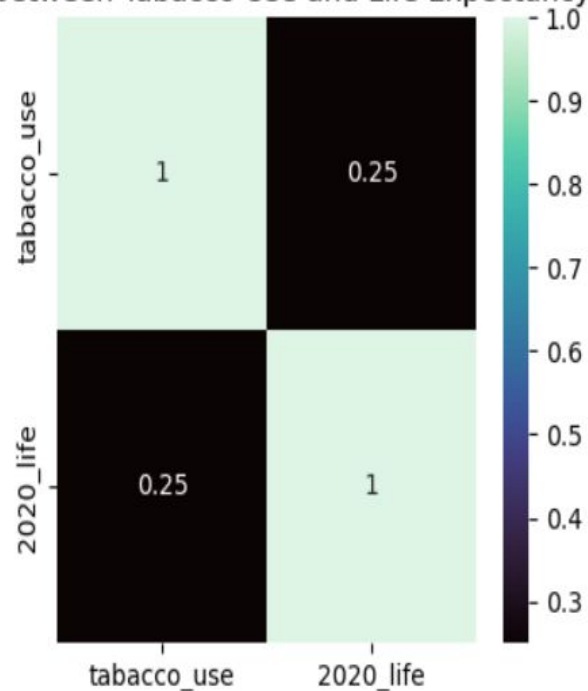
- 1) Performed summary statistics to find values such as mean and standard deviation for each column
- 2) Applied my custom standard deviation function
- 3) Dug into the trends of the data to answer these questions:
 - Which country has the highest and lowest tobacco use for 2020?
 - Highest -> Myanmar, 44.1
 - Lowest -> Ghana, 3.5
 - Which country has the highest and lowest life expectancy for 2020?
 - Highest -> Singapore, 84.9
 - Lowest -> Lesotho, 52.0



Data Visualization

- Heatmap to show the correlation between tobacco usage and life expectancy in 2020
- Shows a weak correlation of 0.25 between the two variables

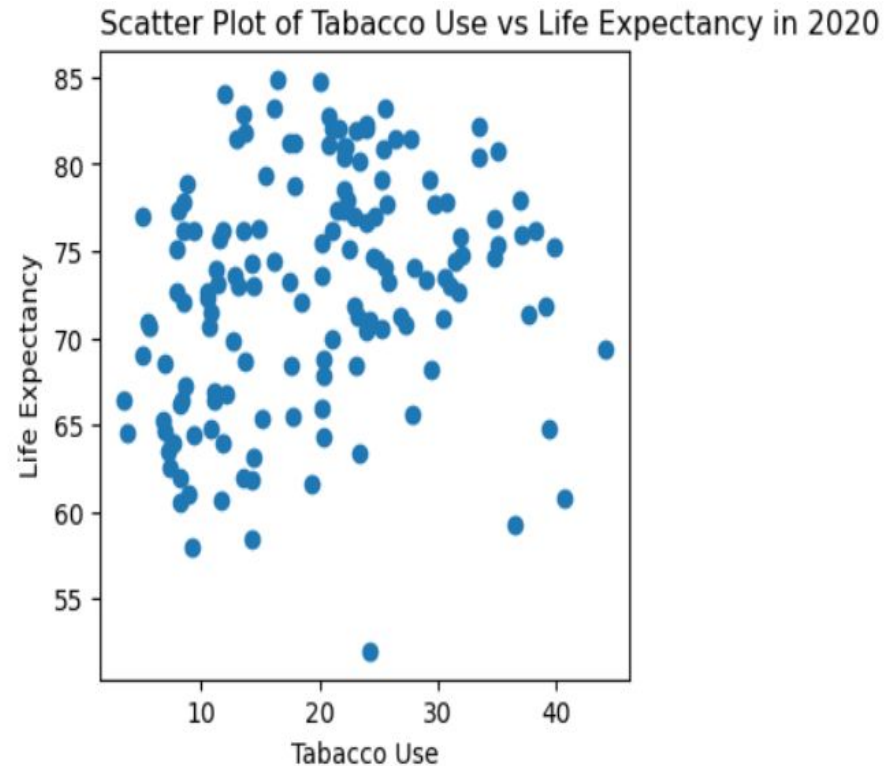
Correlation between Tobacco Use and Life Expectancy in 2020





Data Visualization

- Scatter plot between tobacco usage and life expectancy in 2020
- Shows no relationship between the two variables





Conclusions

- Based on the data shown above, there did not seem to be a relationship between life expectancy and tobacco usage in the year 2020
- The country with the highest tobacco use for 2020 was Myanmar
- The country with the lowest life expectancy for the year 2020 was Lesotho



Recommendations and Side Notes

- Although the data does not show a relationship between tobacco use and a decrease in life expectancy, smoking is still harmful and should be limited or completely stopped
- This data does not take into account how a variety of countries have less access to healthcare or medication, which is a big factor in overall life expectancy
- The data for tobacco usage only considers adults while there could be underage individuals