**Project Proposal**

**Project team details:** Shangfeng Huang, Yash Alpesh Rajpuriya

**Title:** Correlation Analysis of Life Expectancy

**Problem statement.** With the advancement of global healthcare and social development, life expectancy is steadily increasing. Life expectancy serves not only as a crucial indicator of the health status of a country or region but also as a key marker reflecting social development and welfare conditions. Longevity signifies extended periods where individuals can contribute to societal and economic activities, while also indicating the potential for prolonged family life and community engagement. However, life expectancy is subject to a variety of influences, including cultural practices and the region or country of residence. In this project, our aim is to explore the primary factors influencing change in life expectancy, including but is not limited to government healthcare expenditure, government GDP, vaccination rates, body weight, alcohol consumption, educational attainment, and geographical region. By exploring the correlation between life expectancy and these factors, our summary analysis can serve as a reminder for governments and individuals to pay more attention to their lifestyle habits and physical health.

**Objectives.** The database used in this project containing information about Life Expectancy and 18 potential influencing factors, for the years 2000 and 2015, was gathered from 193 countries from world health organization (WHO)’s global health observatory (GHO). Performing Exploratory Data Analysis (EDA) on the dataset and visually analyzing the impact of various factors on life expectancy can provide valuable insights. Observing the results allows us to understand the extent to which each factor influences life expectancy. For instance, we may find that government expenditure on healthcare significantly impacts life expectancy. Furthermore, we aim to explore the combined effects of multiple factors on life expectancy, such as the impact of population size and healthcare expenditure, as well as the influence of per capita income and per capita healthcare expenditure on life expectancy. Finally, based on the provided data, we will construct a predictive model to estimate life expectancy values based on various inputs, such as healthcare expenditure. This model serves as a tool for governments to monitor the life expectancy levels of their population and reminds individuals to pay attention to their personal health and well-being.

**Possible solutions**

**Current project progress**