Life Expectancy Analysis – Data Analytics Project

📌 Project Overview

This project analyzes factors influencing **life expectancy** using data from the **World Health Organization (WHO)**. The goal is to uncover key socio-economic, health, and environmental determinants affecting life expectancy and provide **data-driven recommendations** for policymakers.

🔍 Research Questions

1. What are the key predictors of life expectancy?
2. How do economic factors (GDP, education) impact life expectancy?
3. What is the role of healthcare and immunization in increasing life expectancy?
4. How do malnutrition and diseases like HIV/AIDS influence life expectancy?

📊 Tech Stack Used

- **Programming Language:** Python
- **Libraries:** Pandas, NumPy, Matplotlib, Seaborn, Scikit-learn, Statsmodels
- **Machine Learning Model:** Multiple Linear Regression

🔬 Key Analysis Steps

1. **Data Cleaning & Preprocessing**: Handled missing values, transformed features, and created interaction terms.
2. **Exploratory Data Analysis (EDA)**: Identified correlations, visualized trends, and detected outliers.
3. **Feature Engineering**: Created meaningful features like GDP_Schooling and HIV_AdultMortality.
4. **Model Training & Evaluation**: Built a regression model, optimized features, and evaluated using RMSE & R^2.
5. **Insights & Recommendations**: Suggested actionable steps for improving life expectancy based on data findings.

📈 Results

- **Model Performance**:
  - R^2 Score: **0.71** (71% variance explained)
  - RMSE: **5.20** (average error in predicting life expectancy)
- **Key Findings**:
  - **HIV/AIDS (-4.01)** has the strongest negative impact on life expectancy.
  - **Higher GDP and schooling** positively impact life expectancy.
  - **Malnutrition (thinness 1-19 years)** and **infant mortality rate** reduce life expectancy.

🏆 Impact & Recommendations

- Strengthen **HIV/AIDS prevention & treatment programs**.
- Improve **nutrition & healthcare accessibility** for children and adolescents.
- Invest in **education & economic growth** to enhance public health outcomes.