**Abstract:**

The emergence of COVID-19 in late 2019 led to global lockdowns and various health complications, including increased rates of diabetes. This study analyzes datasets representing the percentage change in diagnosed diabetic individuals across different age groups in various states of the USA before (2018 to 2019) and after (2019 to 2020) the onset of COVID-19. The goal is to visualize these changes and determine if there is any correlation between the increase or decrease in diabetic percentages and the emergence of COVID-19.

**Introduction:**

COVID-19, declared a pandemic in 2020, prompted widespread lockdowns globally. This had significant health implications, including impacts on diabetes rates across different age groups in various states of the USA. Understanding the correlation between COVID-19 and diabetic percentages can provide insights into the broader health effects of the pandemic.

**Problem Statement:**

This project aims to visualize the percentage change in diagnosed diabetic individuals across different age groups in various states of the USA before and after COVID-19. Additionally, it seeks to determine if there is a correlation between the increase or decrease in diabetic percentages and the emergence of COVID-19.

**Methodology:**

**Data Collection and Preprocessing:**

The datasets used in this analysis represent the percentage increase/decrease of diagnosed diabetic individuals across different age groups in various states of the USA.

Each dataset is processed to focus on essential columns, including the Year and Percentage columns for different age groups (18-44, 45-64, 65-74, 75+).

The Year column is converted to datetime format to facilitate temporal analysis.

Missing or invalid data entries are handled by dropping rows with missing values, ensuring data integrity for further analysis.

**Visualization:**

The primary visualization tool used is Matplotlib, a popular Python library for creating static, animated, and interactive visualizations.

Line plots are employed to visualize the percentage change in diagnosed diabetic individuals before (2018 to 2019) and after (2019 to 2020) the onset of COVID-19 for each state and age group.

The plots provide a clear comparison of trends, highlighting any significant changes in diabetic percentages across different demographics and geographic locations.

Correlation Analysis:

Pearson correlation coefficients are computed to quantify the strength and direction of the linear relationship between the percentage changes in diagnosed diabetic individuals before and after COVID-19 for each state and age group.

Correlation coefficients close to +1 indicate a strong positive linear relationship, while coefficients close to -1 indicate a strong negative linear relationship. A coefficient near 0 suggests no linear relationship.

The correlation analysis helps determine if there is a statistically significant correlation between the emergence of COVID-19 and changes in diabetic percentages, providing insights into potential health impacts of the pandemic.

**Results/Analysis:**

The visualizations depict the percentage change in diagnosed diabetic individuals across different age groups in various states before and after COVID-19. The correlation analysis reveals the correlation coefficients between these percentage changes, providing insights into potential relationships between COVID-19 and diabetes rates.

**Conclusion:**

The analysis of diabetic percentage changes before and after COVID-19, along with correlation coefficients, sheds light on the impact of the pandemic on diabetes rates across different demographics and geographic locations. These findings contribute to understanding the broader health implications of COVID-19 beyond its direct effects, highlighting areas for further research and public health interventions.