

PREAPARED BY :

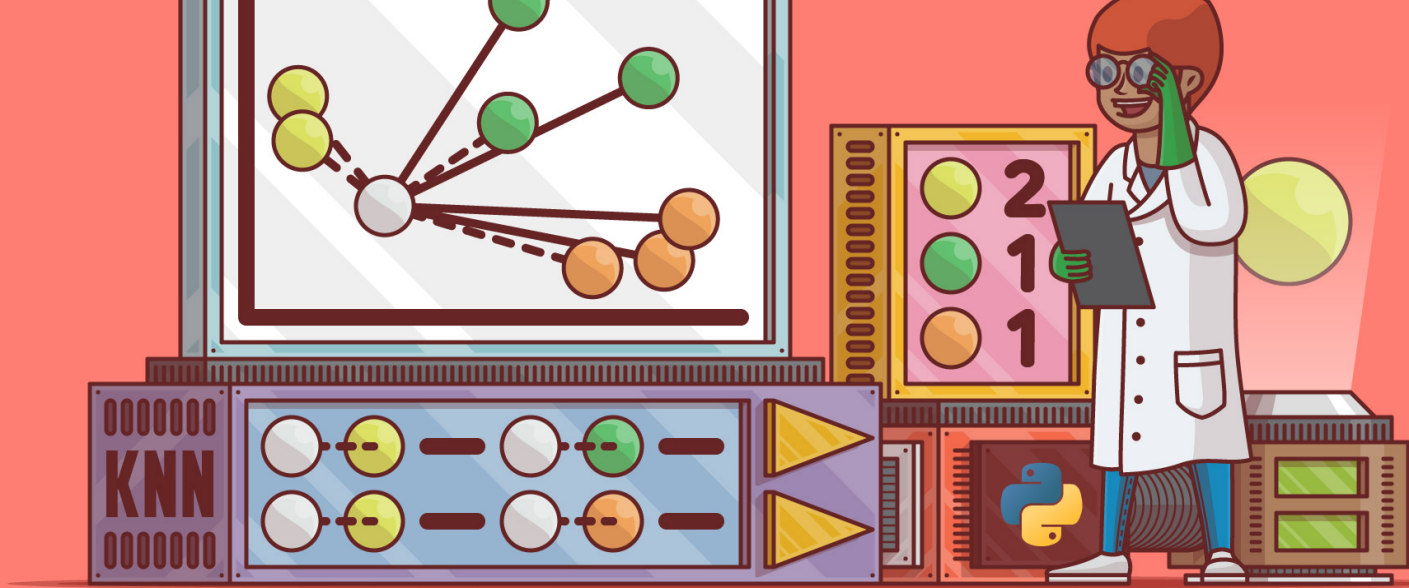
Siphosethu Raga



DOT
GLASSES

**Data
Cleaning**

2023-2024



Overview

In this data cleaning journey, my aim was to ensure the dataset's reliability and suitability for analysis. I started by loading the dataset from an Excel file using Pandas, a powerful tool for handling data in Python. Recognizing the importance of clear column labels, I took the time to simplify them, making the dataset easier to understand. Next, I meticulously checked for missing values, using Pandas to ensure the dataset was complete. Thankfully, no missing values were found, giving me confidence in the dataset's quality.

To gain a better understanding of the dataset's nature, I turned to descriptive statistics. By utilizing Pandas' `describe()` method, I obtained valuable insights into the numerical data's distribution and summary statistics. Additionally, I provided guidance on visualizing the dataset through plotting, setting the stage for insightful exploration. As the sole curator of data integrity and analysis, my efforts resulted in a clean and organized dataset, ready to uncover meaningful insights and inform decision-making processes.