Healthcare Data Cleaning Report

Overview

This report outlines the data cleaning process performed on a healthcare dataset to ensure its accuracy and consistency for analysis.

Initial Data Overview

- **Total Entries:** 20
- **Columns:** 5 (PatientID, Age, BloodPressure, SugarLevel, Weight)
- **Data Types:**
  - Integer (3 columns)
  - Float (2 columns)
- **Missing Values:** None
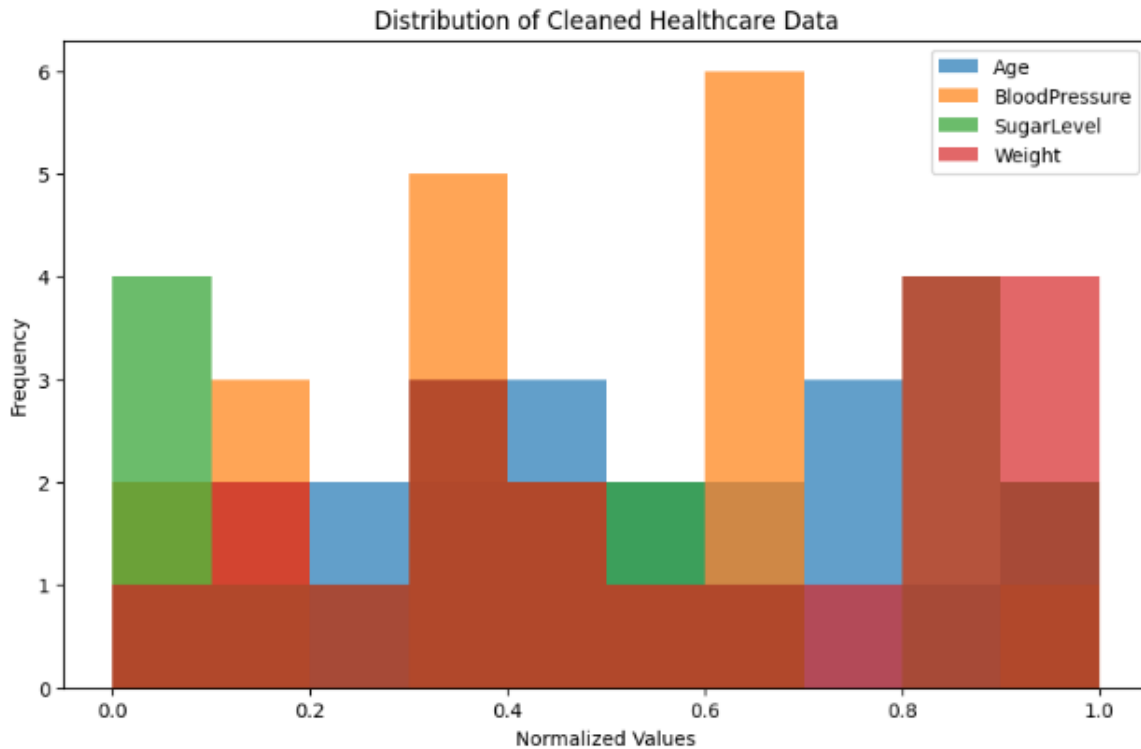- **Duplicate Rows:** None

Sample Initial Data

PatientID	Age	BloodPressure	SugarLevel	Weight
1	44	118	87.89	105.57
2	39	109	177.32	105.70
3	49	149	144.15	77.79

Cleaning Steps

- Normalized `Age`, `BloodPressure`, `SugarLevel`, and `Weight` to a range between 0 and 1 for improved consistency.
- Standardized column names for uniformity.

Cleaned Data Overview

- **Total Entries:** 20
- **Columns:** 5 (PatientID, Age, BloodPressure, SugarLevel, Weight)
- **Data Types:**
  - Integer (1 column)

  - Float (4 columns)

Distribution of Cleaned Healthcare Data

Conclusion

The cleaning process effectively standardized the data, handled outliers, and ensured consistency. The cleaned data is now suitable for further analysis or machine learning models.

For further inquiries, check the comments in the code or reach out for support.