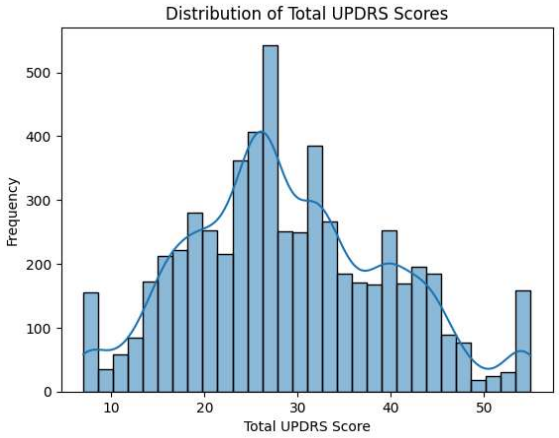
**MILESTONE 1**
**-Harshit Soni (hs5666)**

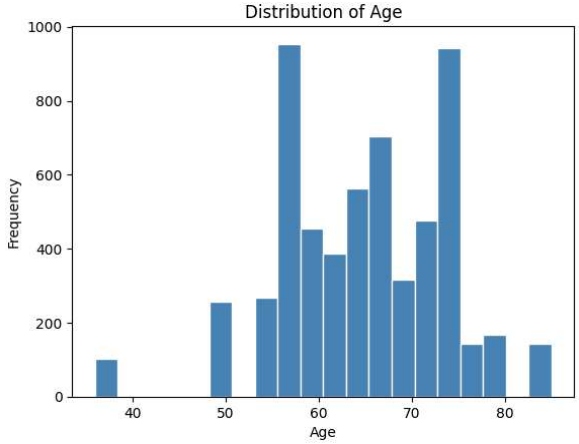
## Motivation:

The goal of this project is to determine whether biomedical voice features can reliably predict the progression of Parkinson's disease as measured by UPDRS scores. By exploring this relationship, the analysis seeks to provide valuable insights that may support non-invasive monitoring techniques in clinical settings. The findings could contribute to enhanced patient care through improved tracking of disease progression and more informed clinical decisions.

## Dataset:

For this project, the Parkinson's Telemonitoring dataset from the UCI Machine Learning Repository will be used. This dataset contains voice recordings from 42 individuals with early-stage Parkinson's disease, totaling about 5,800 instances. Each entry includes a series of biomedical voice measurements and corresponding clinician-assessed UPDRS(Unified Parkinson's Disease Rating Scale) scores, a metric that measures the severity of the disease. The dataset is well-suited for this analysis because it offers comprehensive real-world data that can support rigorous predictive modeling.




| Features | Target Variables |
|---|---|
| 'age', 'test_time', 'Jitter(%)', 'Jitter(Abs)', 'Jitter:RAP', 'Jitter:PPQ5', 'Jitter:DDP', 'Shimmer', 'Shimmer(dB)', 'Shimmer:APQ3', 'Shimmer:APQ5', 'Shimmer:APQ11', 'Shimmer:DDA', 'NHR', 'HNR', 'RPDE', 'DFA', 'PPE', 'sex' | 'motor_UPDRS', 'total_UPDRS' |

## Methodology:

The analysis will focus on predicting the progression of Parkinson's based on various factors involving:

- Correlation: will help assess the relationship between various variables.
- PCA: Due to high dimensionality of data, PCA will be used to extract most important features.
- Regression: To predict the UPDRS Score using all the predictors.

The above methodologies are appropriate because they combine thorough data exploration, sound statistical testing, and effective dimensionality reduction to yield reliable and interpretable insights into Parkinson's disease progression.