Voice Gender Dataset Summary

- **Source**: Available on Kaggle, uploaded by **primaryobjects**

- **Data Type**: Tabular dataset in **CSV format**

- **Number of Samples**: Contains **3,168** labeled voice samples

- **Labels**: Each sample is classified as either **male** or **female**

- **Features**: Includes **20 acoustic features** per voice sample:

  - Mean frequency

  - Standard deviation

  - Median

  - First and third quartiles

  - Interquartile range

  - Skewness

  - Kurtosis

  - Spectral entropy

  - Spectral flatness

  - Mode

  - Pitch

  - Centroid frequency

  - Minimum and maximum frequencies

  - Mean of fundamental frequency

  - Spread and median of fundamental frequency

- **Language**: The original recordings are in **English**

- **Purpose**: Designed for **gender recognition/classification** using machine learning

- **Application Areas**:

  - Speech recognition systems

  - Voice-controlled assistants

- o Call center analytics

- o Human-computer interaction

- **Suitable For**:

  - o Classification tasks

  - o Feature analysis

  - o Educational projects

  - o Signal processing and audio analytics research

- **Usage**: Easily loadable into Python using libraries like pandas or sklearn

- **Dataset Size**: Small and manageable, ideal for quick experiments and classroom demonstrations