Week 10

PROGRESS REPORT

1. Word Frequency Analysis

The code calculates word counts for each song's lyrics and analyzes their distribution:

- Creates a lyrics length distribution DataFrame showing word counts per song
- Calculates key statistics:
    - Mean word count: 340.87 words per song
    - Median word count: 294 words
    - Estimated mode: 231.73 words (using Gaussian KDE)

2. Statistical Modeling

The code attempts to fit the word count distribution to statistical models:

- Uses scipy.stats.skewnorm.fit() to fit a skewed normal distribution
- Returns fitted parameters: (204208556.727, 122.999, 276.727)
- This suggests the word count distribution is right-skewed, with most songs having word counts below the mean

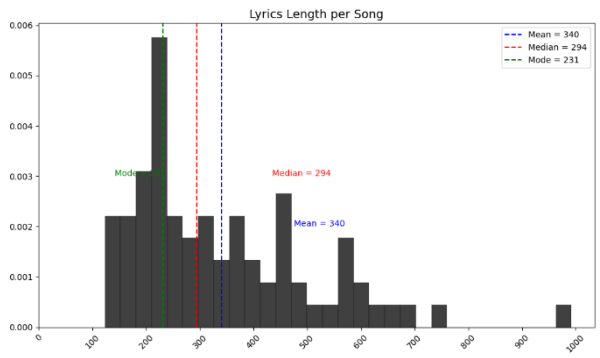
3. Visualization and Reporting

The code includes visualizations and formatted reporting:

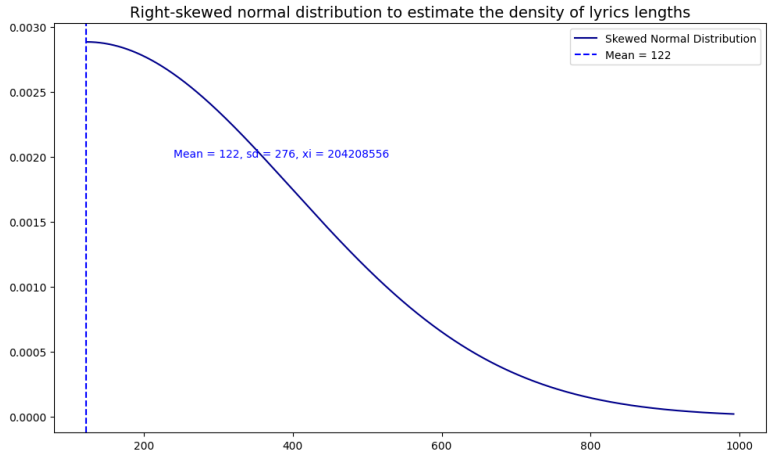
- Displays top 10 songs by word count in an HTML table with color coding
- Shows basic statistics about the dataset:
    - Total songs: 78
    - Unique composers: 17
    - Unique lyricists: 34

We Created 3 plots

1. A histogram for lyrics length per song



2. Right-skewed normal distribution to estimate the density of lyrics lengths



3.